

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2002, 05:48:05 ; Search time 2447.42 Seconds

(without alignments)
6441.250 Million cell updates/sec

Title: US-09-245-198a-1

Perfect score: 1168

Sequence: 1 ggtgctgagcctggcctg.....ataatcatgtctcttc 1168

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
EST :
1: em_estba :
2: em_esthum :
3: em_estin :
4: em_estnu :
5: em_estov :
6: em_estpl :
7: em_estro :
8: em_hic :
9: gb_est1 :
10: gb_est2 :
11: gb_hic :
12: gb_gss :
13: em_gss_hum :
14: em_gss_inv :
15: em_gss_pln :
16: em_gss_vrl :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	883.8	75.7	1033	11	AK020909 Mus muscu
2	625.8	53.6	918	10	BE577781 BE577781
3	519.4	44.5	731	10	BI871711 BI871711
4	507.4	43.4	728	10	BI870393 BI870393
5	504	43.2	561	9	AW763237 AW763237
6	497	42.6	533	10	BE628951 BE628951
7	489.6	41.9	650	10	BE628951 BE628951
8	481.2	41.2	584	9	AW917574 AW917574
9	480.4	41.1	777	10	BI819200 BI819200
10	473	40.5	834	10	BI766766 BI766766
11	456.4	39.1	828	10	BI596681 BI596681
12	437.8	37.5	471	9	AA221610 AA221610
13	390.6	33.4	445	9	AA870722 AA870722
14	360	30.8	360	10	BE654876 BE654876
15	353	30.2	405	9	AI854476 AI854476
16	331	28.3	367	10	BI111534 BI111534
17	325.4	27.9	376	9	AA792068 AA792068

18	309	26.5	474	10	BI965174 BI965174
19	306	26.2	315	10	BF466521 BF466521
20	299.4	25.6	538	10	BF821434 BF821434
21	297.2	25.4	400	9	AI152313 AI152313
22	292.8	25.1	493	10	BE307031 BE307031
23	289.6	24.8	443	10	BF283688 BF283688
24	289.4	24.8	542	10	BF041509 BF041509
25	285.2	24.4	443	10	BF378802 BF378802
26	284.4	24.3	785	10	BI762980 BI762980
27	283.8	24.3	404	10	BF044430 BF044430
28	278.4	23.8	897	10	BI730298 BI730298
29	274.4	23.5	412	9	AI010416 AI010416
30	271	23.2	581	10	BI738634 BI738634
31	270.4	23.2	402	10	BF410871 BF410871
32	270.2	23.1	292	9	BB268794 BB268794
33	258.2	22.1	558	10	BM484863 BM484863
34	255.2	21.8	394	10	BI300370 BI300370
35	243.6	20.9	371	10	BI300370 BI300370
36	242.4	20.8	441	10	BI967060 BI967060
37	242.2	20.7	374	9	AA800970 AA800970
38	240	20.5	240	9	AW764050 AW764050
39	237	20.3	894	10	BI908274 BI908274
40	236.6	20.3	342	9	AA637970 AA637970
41	232.4	19.9	322	10	BI299581 BI299581
42	228.8	19.6	318	10	BI300642 BI300642
43	227.4	19.5	345	9	AA875031 AA875031
44	227.2	19.5	345	10	RS5379 RS5379
45	226.4	19.4	320	9	AI408018 AI408018

ALIGNMENTS

RESULT 1	AK020909	1033 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AK020909				
DEFINITION	Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone: A930030D13; tumor necrosis factor (ligand) superfamily, member 12, full insert sequence.				
ACCESSION	AK020909				
VERSION	AK020909.1	GI:12861640			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain: C57BL/6J) adult retina cDNA to mRNA, clone: lib-RIKEN full-length enriched mouse cDNA library				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2 (sites)				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, T., Muramatsu, M., and Hayashizaki, T.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	2049374				
PUBMED	11042159				
REFERENCE	3 (sites)				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasai, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, Y., Yoneda, Y., Ishikawa, T., Ozawa, Y., Tanaka, T., Matsura, S., Kawai, J., Okazaki, T., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system -84-format sequencing pipeline with 384 multichannel sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				

PUBMED	20530913
JOURNAL	11076861
REFERENCE	4 (sites)
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5 (bases 1 to 1033)
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Atakawa,T., Baladrelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Horii,F., Hume,D., Imclan,K., Ishii,Y., Itoh,M., Izawa,M., Katsukawa,T., Kato,H., Kawaji,K., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schirml,L., Shibata,K., Shibata,Y., Shinaigawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (18-AUG-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GGAAGACAAAGGATCCAGAGCTCTTTTTCCTTTTTCCTTA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through two rounds of normalization to Rot - 20.0 and subtraction to Rot - 458.8. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGGAGATTGTCCAGTTAATTATATGCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B. Retina RNA was provided by Stefano Guslinch (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Retina RNA was provided by Stefano Guslinch (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.
FEATURES	source location/Qualifiers 1..1033 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="MGD:MGI:1911867" /db_xref="taxon:10090" /clone="A930030DI3" /tissue_type="retina" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult"
gene	1..433 /gene="tnsf12"
CDS	<1..433 /gene="tnsf12" /note="data source:MGD, source key:MGI:1196259, evidence:ISS putative tumor necrosis factor (ligand) superfamily, member 12" /codon_start=2 /protein_id="BAB32249.1"

/db_xref="GI:12861641"									
/translation="IAHYHVRPPGDDGAGDGTVSGMEETKINSSPLRDRI									
GEFPIVIRAGILYLKCYHPEDEKNAVYIKLIDLYNGVALKCLEFSAITAASSGFPOLR									
LCQVSGILPLRPSSLEKIRTLPMHLELAAFFLITFGLFQVH"									
BASE COUNT 228 a 300 c 248 g 256 t 1 others									
ORIGIN									
Query Match 75.7%: Score 883.8; DB 11; Length 1033;									
Best Local Similarity 99.48; Pred. No. 6.9e-205;									
Matches 918; Conservative 0; Mismatches 3; Indels 3; Gaps 3;									
QY	248	atcgaccccatcttgaggttcacccctcggccagagacagatggagcacaagcagtg	307						
Db	2	ATTGAGACCCCATTTATGAGGTTTCATCCTCGGCCAGACAGATGAGCACAAGGTGTG	61						
QY	308	gatggagacatgtagtgcctgggaagagacacaatatcaagctccagccctcgcgtac	367						
Db	62	GATGGAGACAGTAGGCGCTGGGGAAGAGACCAAAATCAACACTCCAGGCCCTCGCGTAC	121						
QY	368	gaccgcagaatctggggaattcacagtcacaggcctggcctactaacctgactctag	427						
Db	122	GACCCCGCATTTGGGGAATTTACACTCATCAGGGCTGGGCTCTACTACTGTACTGTAG	181						
QY	428	gtgcacttgaatggaggaagagctgtctactgaagctggaactgtggtgaaagtg	487						
Db	182	GTGCATTTGATGAGGGAAGAGGCTGTCTACCTGTAAAGCTGCACTTGCTGTGAACGGTGTG	241						
QY	488	ctggccctgcgctgccttggaagaattctccacacacagagaagaagctctccggccacg	547						
Db	242	CTGGCCCTGCGCTGCTCTGGAAGAAATCTCAGCCACAGACAGACTCTCTGAGGCCACAG	301						
QY	548	ctccglttgtagcagggtgctcgagctgtgtgcgcgtcgagccagggctctccctcgatc	607						
Db	302	CTCCGTTGTGCGAAGTGTCTGGGCTGTGGCGGTGGCGGACGGATCTTCCCTTCGATTC	361						
QY	608	cgcacccctccctgggctcattctaaagctgcgccctctccaaactatttggactctt	667						
Db	362	CGCACCCCTCCCTGGGCTCATCTTTAAAGGCTCGCCCTTCCTCAACTACTTTGGACTCTTT	421						
QY	668	caagctcaactgaaggagctgtctccacagattccctaaactctcccgagctccagagac	727						
Db	422	CAAGTTCACTAGGGGCTTGCTGTCCCAAGATTCCTTTAACTTTCCCGGCTCAGAGAC	481						
QY	728	atcacacacactccctacccaccccccaactccctccacccccctcgtctccttggtccag	787						
Db	482	ATCACACACACCTMTCTACCCACCCCACTCTCCACCCCTAGCTGTCTCTTGGTCCAG	541						
QY	788	tcctgt-ctctccctaaaggacagcagagctgtgtcacatb-cttccatccacagagct	845						
Db	542	TCTGTCTCTCTCTCAAAAGGACGACGACGCTTGTTCACATGTTTTCATTCCACAGAGCT	601						
QY	846	atcctgtctctc-ctaaatcccatccacacacaaactatccaaactcaactagctccaaa	904						
Db	602	ATTCCTTGTCTTCTTTAATATCCATCCATCCACACCAATATCACTCACTGAGCTCCCAA	661						
QY	905	ggccctacttatcccttgactccccacacacacactaccggaacagtgattatgactgtc	964						
Db	662	GCCTCTACTTTATCCCTGACTCCCCACCCACACTCACCCACACGAGTATTATGACTTTGT	721						
QY	965	gaaccaggaactgagatgagctgtagacctcgtgtagcaggaagccagagaacctgggactagg	1024						
Db	722	GCACGACGACATGAAATGGGCTGGACCTGTGGGACGAGAAAGCAGAAACCTGGGACTAGG	781						
QY	1025	ccagaagttcccaactgtgaagggggaagagctgggagcaagctcctccctggatccctgtc	1084						
Db	782	CCAGAGAGTTCCCAACTGTGAGGGGGGGAAGAGCTGGGGAAGAAGTCTCTCCCTGGATCCCTGT	841						
QY	1085	gaattcttgaagaactatttttatcttatctgtgagcaaaatgtttaaattgatatitaa	1144						
Db	842	GGATTTTGAAGAGATACATTAATTTTATTTATTTATGTGACAAAATGTTAAATGATATTAA	901						
QY	1145	gagaaataacatgattctctctc 1168							

[illegible]

Db	300	TGGCTGGAGAGACCAAAATACAAGCTCCAGCCTTGCCTACGACGCCAGATTGG	359
Oy	382	ggaattcacagcacaacaggctggcgctctaactactgtlactgttcagagtgaccttgalga	441
Db	360	GGAATTTCAGTCATCAGGGCGTGAGCTGTACTACTGTACTGTCAAGTGACATTTGATGA	419
Oy	442	gggaagaagctgtlctaacctgaaagtgtgacttgcgtgtagaaggtgtgtgcctgcgtg	501
Db	420	GGGAAAGGCTGTCTACCTGAAAGCTGGACTGCTGTGTGAACGGTGTGTGGCCCTCGCTG	479
Oy	502	cctggaagaattcttaagccacagcgcaagctctctctgtggccccagcttccgtttgcc	561
Db	480	CCTGGAAGAATTCAGCCACGACGCAAGCTCTCTGGCCCCGACGCTCCGTTTTGCGCA	539
Oy	562	gggtctcgtggcgtgttcgcagctgcagcgaggtcttcccctgtgatccgaaccctccctg	621
Db	540	GGTGTCTGGGCTGTTGGCCGCTCGGGCCA-GGCTTTCCTCTTGGAATCCGACCCCTCCCTG	598
Oy	622	ggcatatcttaagctgtccccctctctaactacttggactcttccaagtactagag	681
Db	539	GGCATATCTTAAGCGCTG-CCCCTTCTTAACATACTTTGGACTCTTCAAGTCACTAGAG	657
Oy	682	ggcct c686	
Db	658	GGCCT 662	
RESULT 3			
B1871711		731 bp	mRNA linear EST 11-OCT-2001
LOCUS	603395825F1 NIH_MGC_90 Homo sapiens CDNA clone IMAGE:5405478 5,		
DEFINITION	mRNA sequence.		
ACCESSION	B1871711		
VERSION	B1871711.1 GI:16045386		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 731)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LHAM12034 Row: b Column: 07 High quality sequence stop: 728.		
FEATURES			
source	Location/Qualifiers		
	1..731		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5405478"		
	/clone_lib="NIH_MGC_90"		
	/library_type="adenocarcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: liver; Vector: pCMV-Sport6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: This is a NIH_MGC Library."		
BASE COUNT	125 a 242 c 228 g 136 t		
ORIGIN			

DB	59	TCACGAGGAGGCGCTGCTCCTCCCG	683
RESULT	5		
LOCUS	AW763237	561 bp	linear
DEFINITION	AW763237	mus musculus	CDNA clone IMAGE:3155633 5'
ACCESSION	AW763237		
VERSION	AW763237.1	GI:7695174	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 561)		
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Other_ESTs: ur70d09.xl Contact: Robert Strausberg, Ph.D. Email: cgaaps-remail.nih.gov Tissue Procurement: lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D. CDNA Library Preparation: Life Technologies, Inc. DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: image.llnl.gov/image/html/lresources.shtml		
FEATURES			
Source	MGI:1058389 Seq primer: -40RP from Gldco High quality sequence stop: 433. Location/Qualifiers 1..561 /organism="Mus musculus" /strain="129 - C57/B6 - EVBN" /db_xref="taxon:10090" /clone="IMAGE:3155633" /clone_lib="NCI-CGAP_Mam3" /tissue_type="tumor, gross tissue" /dev_stage="10 months" /lab_host="DH10B" /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dr. providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."		
BASE COUNT	108 a 158 c 194 g 100 t	1 others	
ORIGIN			
Query Match	43.2%	Score 504;	DB %; Length 561;
Best Local Similarity	99.8%;	Pred. No. 2.6e-112;	
Matches 504;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
1	ggtctgagcttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	60	
2			
3	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	120	
4			
5	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	180	
6			
7	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	240	
8			
9	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	300	
10			
11	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	360	
12			
13	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	420	
14			
15	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	480	
16			
17	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	540	
18			
19	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	600	
20			
21	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	660	
22			
23	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	720	
24			
25	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	780	
26			
27	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	840	
28			
29	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	900	
30			
31	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	960	
32			
33	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1020	
34			
35	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1080	
36			
37	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1140	
38			
39	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1200	
40			
41	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1260	
42			
43	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1320	
44			
45	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1380	
46			
47	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1440	
48			
49	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1500	
50			
51	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1560	
52			
53	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1620	
54			
55	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1680	
56			
57	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1740	
58			
59	ggtctgagccttgagcctgagcctgagcctgagcctgagcctctctgtgtgtgtcagcct	1800	
60			
61	ggtctgagccttg		

Db	237	CTGGGAACACACTAGTCGCGCCCTCGAAGAAAGTCTCTCTTAAGGCGCGAAGGCGGCGCTCG	296
Qy	241	ccgagcataltgacgccataltgaggttcatctctcgcgcagagacagatlgagacaagc	300
Db	297	CCGACCTATTGACGCCCTATTATGAGGTTTCATCTCTGCGCAGACAGGATGAGACACAAGC	356
Qy	301	aggcttgatitggagaaatgagatgctgtggagagaagacccaataacagatccagccctc	360
Db	357	AGGTGTGATGGAGACAGTAGAGTGGCTGGGAAGAGACCAAAATCAACACACTCCAGCCCTCT	416
Qy	361	ggcgacacaccccgagatitgggaattacagtaacagtcagggctgtggcctctactctgta	420
Db	417	GGGCTACACGACCGCCAGATTGGGGATNTACAGTCAATGAGGCGTGGGCTCTACTACTCTTA	476
Qy	421	ctgtcagatgtcaactltgatgaggaagcgtgtctactctgaagctgagatctgtgtgaa	480
Db	477	CTGTAGAGTGTCCTTTGATGAGGGAAAGCCTCTACTGTAAGCGTGAGCTGTGGTGAA	536
Qy	481	cggatgtcgtgcgcctgcgctgcctg	505
Db	537	CGGTGTGCTGGCCCTGCGCTGCTG	561
RESULT	6		
LOCUS	BE628951	533 bp	mRNA
DEFINITION	un30c03.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone IMAGE:3373444 5' similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS ;, mRNA sequence.		
ACCESSION	BE628951		
VERSION	BE628951.1	GI:9911639	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 533)		
AUTHORS	NCI-CCGP http://www.ncbi.nlm.nih.gov/ncicgap .		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (infoimage.llnl.gov) for further information. MGI:1083048		
FEATURES	Seq primer: -40RP from Gldco High quality sequence stop: 437. Location/Qualifiers		
source	1..533		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:3373444"		
	/clone_1lb="Soares_mammary_gland_NMLMG"		
	/sex="female (lactating)"		
	/tissue_type="mammary gland"		
	/lab_host="DH10B"		
	/note="Vector: pRT30-Pac (Pharmacia) with a modified polylinker. 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - Oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."		
BASE COUNT	91 a 206 c 105 g 131 t		
ORIGIN			
Query Match	42.6%:	Score 497:	DB 10: Length 533:
Best Local Similarity	99.4%:	Pred. No. 1.3e-110:	
Matches 530: Conservative	0: Mismatches	0: Indels	3: Gaps 3:

481 cgggtgctgagccctgctgctgctggaagaatctcaagcagcagcagcctctctg 540
|||||
1 cgggtgctgagccctgctgctgctggaagaatctcaagcagcagcagcctctctg 60
541 gccccagcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 600
61 gccccagcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 120
601 tcgagctcagccctcctcctcctcctcctcctcctcctcctcctcctcctcct 660
121 tcgagctcagccctcctcctcctcctcctcctcctcctcctcctcctcctcct 180
661 actctctcaagctcagctgaggggctgctgctcctcagatctcctaaacttccctgctc 720
181 actctctcaagctcagctgaggggctgctgctcctcagatctcctaaacttccctgctc 240
721 caggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780
241 caggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
781 ggtcagctcctgctcctcctcctcctcctcctcctcctcctcctcctcctcct 838
301 ggtcagctcctgctcctcctcctcctcctcctcctcctcctcctcctcctcct 360
839 cagagcagctcctgctcctcctcctcctcctcctcctcctcctcctcctcctcct 897
361 cagagcagctcctgctcctcctcctcctcctcctcctcctcctcctcctcctcct 420
898 tcccaagagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 957
421 tcccaagagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 480
958 actctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1010
481 actctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 533

RESULT 7
BG404836 650 bp mRNA linear EST 12-MAR-2001
LOCUS 602420160F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4527038 5',
DEFINITION mRNA sequence.
ACCESSION BG404836
VERSION BG404836.1 GI:13298284
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 650)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLNL0435 row: h column: 15
High quality sequence stop: 468.
Location/Qualifiers
1..650
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4527038"
/clone_1lb="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1; NCI;"

FEATURES
Source

Site-2: Salt; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."

BASE COUNT 157 a 194 c 149 g 150 t
ORIGIN

Query Match 41.9%; Score 489.6; DB 10; Length 650;
Best Local Similarity 92.0%; Pred. No. 8.9e-109;
Matches 550; Conservative 0; Mismatches 44; Indels 4; Gaps 3;

291 gagcacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 350
25 gacccattcagagtgatgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 84
351 ccagccctctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 410
85 ccagccctctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 144
411 actacctgactgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 470
145 actactgactgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 204
471 tgcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 530
205 tgcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 264
531 gctctcctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 590
265 gctctcctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 323
591 gctctcctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 650
324 gctctcctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 383
651 cctacttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 710
384 cctacttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 443
711 tccctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 770
444 tccctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 503
771 gctctcctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 829
504 gctctcctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 563
830 tccatccacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 887
564 atcat--ccaaagagatgactgactgactgactgactgactgactgactgactgactgac 619

RESULT 8
AW917574
LOCUS 584 bp mRNA linear EST 25-MAY-2000
DEFINITION EST148878 Rat gene index, normalized rat, norvegicus, Bento Soares
Rattus norvegicus cDNA clone R01E49 5' end, mRNA sequence.
ACCESSION AW917574
VERSION AW917574.1 GI:8083328
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 584)
Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R., and Adams, M.D.
Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
Gene Index
Unpublished (1998)
Contact: Lee, NH

Db 374 CTGGTGATGGTGTCTGGCCCTGCGCTGCGAGAGATTCTTCAGCCACTGCGGAGT 433

533 tctctctgcccccaagctcgcgtttgtgcagagtgctggcgtgtctgcgctgcgagcaag 592

434 TCCTCTGGGGCCCGACACTCGGCTCTGCGAGGTGTGTGGGCTTTGGCCCTGGGGCAGAG 493

533 tctctctctgcagatccgacccctccctgggagtcacttaaggctgcctccctcctaacc 652

494 TCCTCTGGAGCTTCTTCAGAGTTCAGTGAAGGGGCGCTGCTCCCGCAGTGTGCCAGGCT 613

534 TACTTTCGAGACTTCTTCAGAGTTCAGTGAAGGGGCGCTGCTCCCGCAGTGTGCCAGGCT 613

713 cctgtgcagagagatcaacacacccctccctacccccacccctccacccctc-g 771

614 GCGGGCTCC-----CCTGACAGAGCTCTGTGGGACCCGGTCCCTGCGCCACCTCAG 668

772 ctgtctctgtgtccagctcgtctctcc--tcaaggcagcagagctgtgtacatgt 829

669 CCGCTCTTGTGCTTCAGAGCTGCGCTCCCTCTAGAGAGTGTGGGCGCTGTTCACGTGT 728

QY 830 tcatctcc 837

Db 729 TTCATTC 736

RESULT 10

LOCUS B1766766 834 bp mRNA linear EST 25-SEP-2001

DEFINITION 60305686F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5', mRNA sequence.

ACCESSION B1766766

VERSION B1766766.1 GI:15758344

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 834)

NIH-MGC http://mgc.ncl.nih.gov/.

Unpublished (1999)

National Institutes of Health, Mammalian Gene Collection (MGC)

Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://lmage.lnl.gov

Plate: LLM11517 row: C column: 18

High quality sequence scop: 772.

Location/Qualifiers

1. 834

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5206217"

/clone_lib="NIH_MGC_122"

/lab_host="DH10B"

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NciI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."

BASE COUNT 154 a 287 c 223 g 170 t

ORIGIN

Query Match 40.5%; Score 473; DB 10; Length 834;

Best Local Similarity 78.0%; Pred. No. 1,1e-104;

Matches 653; Conservative 0; Mismatches 160; Indels 24; Gaps 6;

QY 96 ctcaaggagagctgaacagcaagagacgcgcggagagcccccgtgaactgaatcccaagacag 155

1 CCCAGAGAGAGCTGTGTGACAGAGAGACAGACCCCTCGGAATCAATCCAGACAG 60

Db 156 aggaagacgaagatgtgtgtactcttcttggaacaactatgtccgctctgaagaagtgtc 215

61 AAGAAAGCAGAGATCTCGCCCTTCTGTAACCCGACTAGTTGGCTCTGCGAAGTGTAC 120

QY 216 ctaaggccggaagacgcgcgcctgcgcagctatgtgagccatgatgtatgtctc 275

121 CTAAGAGCCGGAACACAGGCTCGAAGAGGATGCGACGCCATTATTAAGTTCAATCAC 180

Db 276 ggc-aggacagatgagacaaagagtgatgagatgagagatgtgtgtggaagag 334

181 GACCTGAGACAGAGAGAGAGGAGGAGGAGTGTGACGGAGAGATGAGTGTGGAGGAA 240

QY 335 accaaatcaaacagctccagccctctgcgtacgacgcgcagatctgggaaattacagt 394

241 GCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAAACCCGAGATCGGGAGTTTATATC 300

QY 395 atcaaggcttgggctctactactgtactgtacagtgtgcaacttgatgagaaagctgtc 454

301 ACCCGGCTGGGCTCTACTACTCTGTACTGTCAAGTGCACTTTGATGAGGAGAGGCTGTC 360

QY 455 taactgaagctgtgactgtgtgaaagtggtgtgcctgcgcctgcgtcgtgaagaatlc 514

361 TACTGAAAGCTGTGACTTCTGTGTGATGTGTGTGCTGCGCCCTGCTGCGTGAAGAAATTC 420

QY 515 tcaagcaagcagcagagctctcctgtggcccaagctccgtttgtgcaagtgtctgtgtgtc 574

421 TCAGCCACTGTGGCGGCGAGTTCCTCGGGGCCCAAGCTCCCTCTGCGAGGTGTGTGGCTG 480

QY 575 ttgcgcgtgcgcgcagaggtctctcctcttgatgcgcacccctccctgtgcacatcag 634

481 TTGGGCTGTGGGCA-6GTCTCCTCTGTGGGATCCGACACCTCCCTGGGCCATCTCAAG 539

QY 635 gctgcctccctcttaacttaacttgaactcttcaagttcaactgaagggtgcctgtcctcc 694

540 GCTGCCCTTCTCTCACTTCTGAGACTTTCAGGTTCTGAGTGTGAGGGGCTCGGTCTC- 598

QY 695 cagatctcttaacttccctgtgtcctcagagagatcacacaactccct----accacac 750

599 -----CCACAGTGTGTCCAGGCTGCGGGCTCCCTGACAGCTCTGTGGGACCCGGT 652

QY 751 ccccaactcccaacccctgcgtctcctctgtgtcagtcgtgtctcct--tcaaggca 808

Db 653 CCCCTTGTGCCCAACCTCAGCGCGCTCTTGTCTCAGACCGGCGCTCTCTAGAGGCT 712

QY 809 gccagagctgtgtcacatgttccatcc-----acagagctatcctgtcctctc 859

713 GCGTGGGCTGTTCACAGTGTTCATCCACATTAATATAGTATTCACACTTATCTT 772

QY 860 taacttccatcccaacaaatacactactactagcttccaaagccctactat 916

773 AAAAATTCCTCCCAAGCGCCACTCTCAGCTTATTAAGTCCCATTTCCCTGACTTT 829

RESULT 11

LOCUS B1596681 828 bp mRNA linear EST 07-SEP-2001

DEFINITION 603243254F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892 5', mRNA sequence.

ACCESSION B1596681

VERSION B1596681.1 GI:15489620

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 828)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1M11722 row: k column: 13
High quality sequence stop: 776.
Location/Qualifiers

FEATURES

source
1..828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5285892"
/clone_id="NIH-MGC-96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to 10⁶ 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT 155 a 278 c 223 g 172 t
ORIGIN

Query Match 39.1%; Score 456.4; DB 10; Length 828;
Best Local Similarity 82.6%; Pred. No. 1.2e-100;
Matches 550; Conservative 0; Mismatches 106; Indels 10; Gaps 2;
QY 137 gaactgaatccccagagagagagagatgtgtactcttcttgaaacattgtc 196
DB 44 GAATCTGAATCCCGACAGAGAAAGCCAGATCTCGCTTCTGAACCGACTAGTT 103
QY 197 cggcctcgaagaagtgtctcctaagaagcgcgccctcgcgagctatctgacgc 256
DB 104 CGGCTCGGCAAGTGCACCTAAAGGCGGAAACACGGGCTCGAAGAGCGATCGACCC 163
QY 257 cattaatgagttcactcctcgcgagagagatgagagcaagaagtgtgagtggaca 316
DB 164 CATATATGAAGTTTCATCCACGACCTGACAGACGACGACGACGAGTGTCGACGGGACA 223
QY 317 gttagtgctgctgagagagagcaaaatcaaacgctcgcagcctctgagctacagccagc 376
DB 224 TTGAGTGGCTGGAGAGCCAGATACACAGCTCCACCTCTGCGCTACACCGCCAG 283
QY 377 attgggaatttaacagtcacatcaggctggtgctcactacactgtcagagtgcaattt 436
DB 284 ATCGGGGAGTTTATAGTACACCGGGCTGGCTTACTACTACTGCTACTGACAGTTC 343
QY 437 gatgagggaaagctgtctactctgaagctgactgtctgtgaaagtggtgtgacctg 496
DB 344 GATAGGGGAGAGCTGTCTACTCTGAAGCTGACCTGCGTGGTGGATGGTGTCTGCGCTG 403
QY 497 cgtgctcgtgagaaattctcagccacagcagcagcagctcctctgggcccccaatcggtt 556
DB 404 CGCTGCTGGAGAGATTCCTAGCCACTGCGGCAAGTTCCTCGGGCCCCCAGCTCGGCTC 463
QY 557 tgcgaagtgtctggtgtgctgctgctgctgagaggtcttctccttcgagatcgcaccc 616

DB 464 TGCCAGAGTGTGGAGCTGTGGCCCTGCGGAGGGTCTCCGAGTCGCGGATCGGACCCCTC 523
QY 617 ccttgagctatcttaagctgtccccctctcttaacctacttgaacttcaagtctac 676
DB 524 CCGTGGGCGCATTCAGAGGCTGGCCCTCTCTCACTTCTGAGACTTTCACAGTTTCAC 583
QY 677 tgaaggagctgtgctctccagagatctcctaacttccctgtgctcagagatcacaca 736
DB 584 TGAGGGGCGCTGTCTCCCGC-----CACTCTCCAGGCTCGGCGCTCCCTCGACA 636
QY 737 cctccct---accacaccactctccaccctcctcgtctgtctgtgtccagctcgt 793
DB 637 GCTCTGTGGACCGCGGTCCTCTGTGGCCACCCCTGACGACCATCTTTCCTCAGACCTGC 696
QY 794 cctcc 799
DB 697 CCTCTC 702

RESULT 12

AA221610 471 bp mRNA linear EST 13-FEB-1997
LOCUS
DEFINITION
my18409.t1 Barstead mouse heart MRLB3 Mus musculus cDNA clone
IMAGE:696209 5', mRNA sequence.
AA221610
AA221610.1 GI:1840863
VERSION
KEYWORDS
EST.
SOURCE
house mouse.
MUS MUSCULUS
ORGANISM

REFERENCE

1 (bases 1 to 471)
Marrin, S., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Gelsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Materston, R.
The WashU-HMI Mouse EST Project

JOURNAL

Unpublished (1996)
Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:429769
Seq primer: -28m13 rev2 EF from Amersham
High quality sequence stop: 301.
Location/Qualifiers

FEATURES

source
1..471
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="IMAGE:696209"
/clone_id="Barstead mouse heart MRLB3"
/sex="mixed"
/tissue_type="heart"
/dev_stage="6 weeks"
/lab_host="DH10B"
/note="Organ: heart; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTCAAGGAGGAGGCGGCGGCTTTTGTGTGTGTGTGTGTGTGTGT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CTTGAATTCGGTACC], digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."
BASE COUNT 91 a 141 c 128 g 111 t
ORIGIN


```

Db      285  |||||
           CACTAGCTCCAAAGCCCTACTATCCCTGACTCCCCACCACACTCACCCGACCCTG 226
Oy      951  |||||
           ttatltgacttlygaccaggcaactgagatggctgacctgtygcaggaaagccagag 1010
Db      225  |||||
           TTTATTGACTTTGTGCACACGACACTGAGATGGCTGGAACCTGTGTCAGGAGCCAGAG 166
Oy      1011 |||||
           aacctggagactlagccagaagttcccaactglaagggggaagagctggggacaagctcct 1070
Db      165  |||||
           AACCTGGGACTAGGCGCAGAGTTCCCACTGTGAGGGGGAAGAGCTGGGGACAAAGCTCCT 106
Oy      1071 |||||
           cccctggatccctlyggaatttgaagaataactatlttataatctatctgacacaaatgtt 1130
Db      105  |||||
           CCTGTGATCCTCTGTGATTTGAAAAGATACTATTTTATTATTTATTTGTGACAAAAATGTT 46
Oy      1131 |||||
           aaatgatattcaagagaataaataatcatga 1159
Db      45   |||||
           AAATGGATTTTAAAGAGAAATTAATCATGA 17

```

Search completed: June 22, 2002, 10:48:59
 Job time: 18054 sec

04-SEP-1998; 98MO-US18393.
05-SEP-1997; 97US-0924634.
(UNIV) UNIV WASHINGTON.
Chaudhary PM.
WPI; 1999-205191/17.
P-PSDB; AAW935591.
New Tumor Necrosis Factor family receptor polypeptides and ligands -
useful for diagnosis and treatment of prostate cancer and
developmental or gestational abnormalities
Example VII; Fig 13B; 156pp; English.
This invention describes isolated Tumor Necrosis Factor (TNF) family
receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
their active fragments. APO4 is useful for diagnosing prostate cancer
by determining levels of APO4 in an individual. Prostate cancer can also
be treated using APO4 selective binding agents linked to a therapeutic
moiety. APO4 polypeptides are also useful for identifying selective
binding agents, useful in diagnosis/treatment of disease by binding of
agents to the polypeptide/active fragment which is extracellular, or
expressed on the cell surface. The binding is preferably performed in
vivo. APO4 polypeptides/ active fragments are also useful for screening
for agonists and antagonists by binding and observing the change in APO4
activity. Effective pharmacological agents useful in diagnosis or
treatment of disease are also identified using APO4 polypeptides/active
fragments and APO4 signal transducer molecules that specifically interact
with a cytoplasmic domain of APO4 and detecting a change in level of APO4
activity. The method is performed in vivo or in vitro. APO polypeptides
are all useful as immunogens for preparing antibodies. APO4 is also
useful for diagnosis/treatment of developmental or gestational
abnormalities. APO8 was transfected to human breast carcinoma cell line
MCF-7, and induced apoptosis.

[illegible]

Db	361	gggctctctactcctgtactgtcaggtgcactttgataaggaagctgtctacttgaag	420
Qy	464	ctgagctctgcgtgtgaacagctgtctgctgcctgcgtcgtcgtgaagaattctcagccaca	523
Db	421	ctgagctctgcgtgtgaacagctgtctgctgcctgcgtcgtcgtgaagaattctcagccaca	480
Qy	524	gcagcaagctctcctctgggccccagctccgttctgtccagatgtctggtctgtgcgctg	583
Db	481	gcagcaagctctcctcctgggccccagctccgttctgtccagatgtctggtctgtgcgctg	540
Qy	584	cggccagagctcttcctcccttcggaatccgaacccctccctcggggtcttaattgaagctgcgcccc	643
Db	541	cggccagagctcttcctcccttcggaatccgaacccctccctcggggtcttaattgaagctgcgcccc	600
Qy	644	ttcctaactcacttctgagctcttccaagttaactgaaggggctctgcctccagattcct	703
Db	601	ttcctaactcacttctgagctcttccaagttaactgaaggggctctgcctccagattcct	660
Qy	704	taaaccttcctctgctctccagagcatcaccaacactctcccta	744
Db	661	taaaccttcctctgctctccagagcatcaccaacactctcccta	701
RESULT 3			
AAA49717			
ID	AAA49717	standard; cDNA; 1353 BP.	
XX			
AC	AAA49717;		
XX			
DT	25-SEP-2000	(first entry)	
XX			
DE	Human PRO207 cDNA clone DNA30879-1152.		
XX			
KW	PRO207: human: antitumour; tumour; therapy: cytostatic;		
KW	breast cancer; ovarian cancer; renal cancer; colorectal cancer;		
KW	uterine cancer; prostate cancer; lung cancer; bladder cancer;		
KW	central nervous system cancer; melanoma; leukemia; neoplasm; ss.		
XX			
OS	Homo sapiens.		
XX			
EH	Key	Location/Qualifiers	
FT	CDS	58..807	
FT		/*tag= a	
FT	sig_peptide	58..177	
FT		/*tag= b	
FT	mat_peptide	178..804	
FT		/*tag= c	
XX			
PN	WO200037638-A2.		
PD	29-JUN-2000.		
XX			
PF	02-DEC-1999;	99WO-US28565.	
XX			
PR	22-DEC-1998;	98US-0113296.	
PR	08-MAR-1999;	99WO-US05028.	
PR	21-APR-1999;	99US-0130232.	
PR	28-APR-1999;	99US-0131445.	
PR	14-MAY-1999;	99US-0134287.	
PR	20-JUL-1999;	99US-0144758.	
PR	26-JUL-1999;	99US-0145698.	
PR	15-SEP-1999;	99WO-US21090.	
PR	15-SEP-1999;	99WO-US21547.	
PA	(GETH) GENENTECH INC.		
XX			
PI	Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Masters SA;		
PI	Napier MA, Pitti RM, Wood WI;		
XX			
DR	WPI. 2000-442668/38.		
DR	P-PSDB; AAY95338.		
PT	Novel composition to inhibit neoplastic cell growth or for treating		

PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities

PS Example VII: Fig 13A: 156bp: English.

CC This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC molecule. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the changer in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.

XX Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;

Query Match 44.8%; Score 522.8; DB 20; Length 1030;
 Best Local Similarity 81.6%; Pred. No. 9e-132;
 Matches 643; Conservative 0; Mismatches 137; Indels 8; Gaps 3;

QY 53 gtaagctgagagagctggagcaagctgtgtccagagacgtcttcagagagagctgaca 112
 DB 1 gtaagctgagagagctggagcaagctgtgtccagagacgtcttcagagagagctg 60
 QY 113 gcaag 172
 DB 61 gcaag 120
 QY 173 gtaagctgag 232
 DB 121 gtaagctgag 180
 QY 233 gcaagctgag 292
 DB 181 gcaagctgag 240
 QY 293 gcaagctgag 352
 DB 241 gcaagctgag 300
 QY 353 gcaagctgag 412
 DB 301 gcaagctgag 360
 QY 413 gcaagctgag 472
 DB 361 gcaagctgag 420
 QY 473 gcaagctgag 532
 DB 421 gcaagctgag 480
 QY 533 gcaagctgag 592
 DB 481 gcaagctgag 540
 QY 593 gcaagctgag 652

DB 541 tcaagctgag 600
 QY 653 gcaagctgag 712
 DB 601 gcaagctgag 660
 QY 713 gcaagctgag 771
 DB 661 gcaagctgag 715
 QY 772 gcaagctgag 829
 DB 716 gcaagctgag 775
 QY 830 tcaagctgag
 DB 776 tcaagctgag

RESULT 9

AA0303964 standard; DNA; 898 BP.

AA0303964;

26-SEP-2001 (first entry)

Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.

TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;
 ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
 retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
 rheumatoid arthritis; macular degeneration; arthritis; rheumatism; ds;
 corneal graft neovascularisation; psoriasis; metastatic condition;
 malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
 preneoplastic condition; myocardial angiogenesis; wound granulation;
 scleroderma; vascular adhesion; telangiectasia; ischaemia; human;
 atherosclerotic plaque neovascularisation; coronary atherosclerosis;
 peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;
 fusion protein.

Homo sapiens.
 Synthetic.

Location/Qualifiers
 52..873
 /tag- a
 /product= "Fusion protein comprising a growth hormone
 leader, a leucine zipper multimerisation
 domain,"
 domain"

Key CDS

WO200145730-A2.
 28-JUN-2001.

19-DEC-2000; 2000MO-US34755.

20-DEC-1999; 99US-0172878.

10-MAY-2000; 2000US-0203347.

(IMNV) IMMUNEX CORP.

Wiley SR.

WPI; 2001-417975/44.

P-PSDB; AA003499.

Modulating angiogenesis in a mammal for treating diseases mediated by

angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or

peripheral tissue, by administering antagonist or agonist of TWEAK

receptor

Db 207 tggattttaaacaagatatattttnattatgtgcaaaatgtgtaaatgga 266
 Oy 1138 tattaagaagataaa 1153
 Db 267 tattaagaagataaa 282

RESULT 11

AAT94101
 ID AAT94101 standard; DNA: 53526 BP.

AC AAT94101;

DT 01-JUN-1998 (first entry)

DE Human PKD1 gene.

KW Human; polycystic kidney disease 1; PKD1; treatment; autosomal dominant polycystic kidney disease; APKD; ss.

XX Homo sapiens.

PN WO9744457-A1.

PD 27-NOV-1997.

PF 22-MAY-1997; 97MO-US08799.

PR 03-JUN-1996; 96US-0658136.

PR 24-MAY-1996; 96US-0655360.

XX (GEN2) GENZYME CORP.

PI Burn T, Connors T, Dackowski W, Germino G, Klinger K;

PI Qian F;

PT WPI: 1998-018511/02.

PT Human polycystic kidney disease gene, PKD1 - useful to treat and

PT diagnose human autosomal or adult onset polycystic kidney disease

PS Claim 2; Pages 90-118; 257pp; English.

XX The present sequence is the human polycystic kidney disease 1

CC (PKD1) gene. The PKD1 gene or polypeptide may be used to treat

CC autosomal dominant polycystic kidney disease (APKD), and identify

CC carriers of mutant PKD1 genes, i.e. subjects susceptible to APKD.

CC Antibodies (Ab) that distinguish between normal and mutant PKD1

CC sequences can also be used in diagnostic tests. Anti-PKD1 Ab may

CC also be used to perform subcellular and histochemical localization

CC studies, and to block the function of PKD1. Ab are also useful in

CC rational drug design studies to identify and test inhibitors of

CC PKD1. Sense and antisense sequences derived from the PKD1 gene may

CC used for detection and therapy.

CC Sequence 53526 BP; 8486 A; 17665 C; 15768 G; 11607 T; 0 other;

SO

Query Match 4.0%; Score 46.6; DB 19; Length 53526;
 Best Local Similarity 48.7%; Pred. No. 0.11;
 Matches 127; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

Oy 685 ctgtctctccagattcctaacttccctgtgtccagagatcaccacactcccta 744

Db 34369 cctccctccctaccctccctccctccctccctccctccctccctccctcc 34428

Oy 745 cccacccacactcctccacccctcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 804

Db 34429 cgtctgaagccctccacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 34488

Oy 805 ggcagcagagcgt 864

Db 34489 cctccctccctccctccctccctccctccctccctccctccctccctcc 34548
 Oy 865 tccatccac 924
 Db 34549 ttcctccctccctccctccctccctccctccctccctccctccctcc 34608
 Oy 925 cccacccac 945
 Db 34609 tccctccctccctccctccctccctccctccctccctccctccctcc 34629

RESULT 12

AAT18551
 ID AAT18551 standard; DNA: 53577 BP.

AC AAT18551;

DT 06-MAY-1997 (first entry)

DE Human polycystic kidney disease normal PKD1 gene.

KW Adult onset polycystic kidney disease; APKD; autosomal dominant;

KW mutant; transversion; transition; deletion; insertion; ds.

XX Homo sapiens.

OS

FT Key

FT misc_feature

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

FT old_sequence

Claim 1; Fig 1; 65pp; English.

The present sequence is that of the normal human PKD1 gene from

CC chromosome 16. Mutations in this gene (e.g. transitions,

CC transversions, deletions and/or insertions) are associated with

CC adult-onset polycystic kidney disease (APKD). The PKD1 locus is

CC GC-rich (62.4%). Comparison of this sequence with a previously

CC reported partial cDNA sequence revealed differences at three

CC locations (see features table). The most significant difference is

at position 4566 of the previously reported sequence. The insertion

Db 627 NNCNNCCNCNCNCNCNCNCNCCTCTCNCNCNCNCNCNCNCNCNCNCNCNCNCNCNC 568
Qy 809 gccagagctgttcaatgttccattccacagacgtatccttgccttcttcttcttcttcttcc 868
Db 567 CNC 508
Qy 869 atcccaaccagaactatccacactcactagcttcccaaaagccctacttattccctgaactccc 928
Db 507 CGCCNC 448
Qy 929 caccacatcaccggac 944
Db 447 CCCCCCNC 432

Search completed: June 22, 2002, 11:53:24
Job time: 6741 sec

alignment_scores:

Quality: 1162.00 Length: 225
Ratio: 5.164 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-245-198a-2 x AAV18599 ..

Align seg 1/1 to: AAV18599 from: 1 to: 1168

1 ValLeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVa 17
2 GTGCTAGAGCTGGGCTGGCGCTGGCGCTGCTGGCTGCTGCTGCT 51
17 ValSerLeuGlySerTrpAlaThrLeuSerAlaGlnGlnProSerGln 34
52 GGTCAAGCTGGGGAGCTGGGCACAGCTGTCTGCCCAGGAGCTTCTCAG 101
34 LuGluLeuThrAlaGluAspArgArgGluProProGluLeuAsnProGln 50
102 AGGAGCTGACAGACAGACCGCGCGAGAGCCCTGACTGAAATCCCAAG 151
51 ThrGluGluSerGlnAspValValProPheLeuGluGlnLeuValArgP 67
152 ACAGAGAAAGCCAGAGATGTGTGCTTCTTGGAACAACACTAGTCCGGCC 201
67 oArGArgSerAlaProLysGlyValArgLysAlaArgProArgArgAlaIle 84
202 TCGAGAAAGTCTCTCTAAAGCCCGGAGAGCGCGCTCCCGAGCTATTTG 251
84 lAlAlaHsTrpGluValHisProArgProGlnAspGlyAlaGlnAla 100
252 CAGCCCATATGAGGTTCATCCCGGCCAGGACAGATGAGACACAAGCA 301
101 GlyValAspGlyThrValSerGlyTTPGluGluThrLysIleAsnSer 117
302 GGGTGGATGGGACAGTGTGTGGTGGGAAGACCAAAATCAACAGCTC 351
117 rSerProLeuArgTrpAspArgGlnIleGlyGluPheThrValIleArg 134
352 CAGCCCTCTGGCTACGACCGCCAGATTTGGGAAATTTACAGTATCAGG 401
134 lAgLyLeuTrpTrpLeuTrpCysGlnValHisPheAspGluGlyVal 150
402 CAGGCTCTACTACTGTACTGTACAGTGCACCTTGTATGAGGAAAGGCT 451
151 ValTrpLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArg 167
452 GTCTACCTGAGCTGAGCTGTGTGTAACGGTGTCTGGCCCTGGCGTG 501
167 sLeuGluGluPheSerAlaThrAlaIleSerSerProGlyProGlnLeu 184
502 CCGGAAGAATTTCTCAGCCACAGCAAGAGCTCTCTGGGCCCCACTGCC 551
184 rGluCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeu 200
552 GTTGTGCCAGAGGTCTGGGCTGTCTGCCCTGGCGCCAGAGGTCTCCCTT 601
201 ArgIleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeu 217
602 CGGATCCGAGACCTCCCTGGGCTCATTTAAAGCTGCCCCCTTCTTAC 651
217 rTrpPheGlyLeuPheGlnValHis 225
652 CTACTTTGAGACTTTCAAGTTTAC 676

seq_name: /SID51/gcdata/geneseq/geneseqn-emb1/NA1999.DAT:AA23425

seq_documentation_block:

ID AAX23425 standard; DNA: 701 BP.
XX
AA23425:

XX 18-JUN-1999 (first entry)
DT
XX
DE Mouse TNRL3 DNA.
XX
KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW developmental abnormality; gestational abnormality; prostate cancer;
KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW apoptosis; mouse; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FH CDS 1..636
FT /*tag= a
FT /product= "TNRL3"
XX
XX WO9911791-A2.
XX
XX 11-MAR-1999.
XX
XX 04-SEP-1998; 98MO-US18393.
XX
XX 05-SEP-1997; 97US-0924634.
XX
XX (UNITW) UNITW WASHINGTON.
XX
XX Chaudhary PM;
XX
XX WPI: 1999-205191/17.
XX
XX P-PSDB; AAW93591.
XX
XX New Tumour Necrosis Factor family receptor polypeptides and ligands -
XX useful for diagnosis and treatment of prostate cancer and
XX developmental or gestational abnormalities
XX
XX Example VII; Fig 13B; 156pp; English.
XX
XX This invention describes isolated Tumor Necrosis Factor (TNF) family
XX receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
XX fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
XX their active fragments. APO4 is useful for diagnosing prostate cancer
XX by determining levels of APO4 in an individual. Prostate cancer can also
XX be treated using APO4 selective binding agents linked to a therapeutic
XX moiety. APO4 polypeptides are also useful for identifying selective
XX binding agents, useful in diagnosis/treatment of disease by binding of
XX agents to the polypeptide/active fragment which is extracellular, or
XX expressed on the cell surface. The binding is preferably performed in
XX vivo. APO4 polypeptides/active fragments are also useful for screening
XX for agonists and antagonists by binding and observing the change in APO4
XX activity. Effective pharmacological agents useful in diagnosis or
XX treatment of disease are also identified using APO4 polypeptides/active
XX fragments and APO4 signal transducer molecules that specifically interact
XX with a cytoplasmic domain of APO4 and detecting a change in level of APO4
XX activity. The method is performed in vivo or in vitro. APO polypeptides
XX are all useful as immunogens for preparing antibodies. APO4 is also
XX useful for diagnosis/treatment of developmental or gestational
XX abnormalities. APO8 was transfected to human breast carcinoma cell line
XX MCF-7, and induced apoptosis.
XX
XX Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

alignment_scores:

Quality: 1089.00 Length: 211
Ratio: 5.186 Gaps: 0
Percent Similarity: 99.526 Percent Identity: 99.526

alignment_block:

US-09-245-198a-2 x AAX23425 ..

Align seg 1/1 to: AAX23425 from: 1 to: 701

seq_name: /SID1/gcseqdata/geneseq/geneseqn-emb1/NA1998.DAT.AA47613	seq_documentation_block:
15 LeuValValValSerLeuGlySerTyrProIleuSerAlaGluGluPro 31	15 LeuValValValSerLeuGlySerTyrProIleuSerAlaGluGluPro 31
1 CTGGTCTGGTCACTCTGGGGAGCTGGGCAACGCTGTCTGCCACGAGGCC 50	1 CTGGTCTGGTCACTCTGGGGAGCTGGGCAACGCTGTCTGCCACGAGGCC 50
31 oSerGlnGluGluLeuThrAlaGluAspArgGluProProGluLeuA 48	31 oSerGlnGluGluLeuThrAlaGluAspArgGluProProGluLeuA 48
51 TTCTCAGGAGAGCTGCACAGCAGAGGACCGCGGGAGCCCCCTGAACCTA 100	51 TTCTCAGGAGAGCTGCACAGCAGAGGACCGCGGGAGCCCCCTGAACCTA 100
48 snProGlnThrGluGluSerGlnAspValValProPheLeuGluGlnLeu 64	48 snProGlnThrGluGluSerGlnAspValValProPheLeuGluGlnLeu 64
101 ATCCCAAGACAGAGAAAGCAGAGATGTGGTAACCTTTCTTGGAAACACTA 150	101 ATCCCAAGACAGAGAAAGCAGAGATGTGGTAACCTTTCTTGGAAACACTA 150
65 ValArgProArgArgSerAlaProLysGlyTyrGlyValAspArgArg 81	65 ValArgProArgArgSerAlaProLysGlyTyrGlyValAspArgArg 81
151 GTCCGGCCTTGAAAGACTGCTCTTAAGGCCGGGAAGGCCGGCCTCGCGG 200	151 GTCCGGCCTTGAAAGACTGCTCTTAAGGCCGGGAAGGCCGGCCTCGCGG 200
81 galAlaIleAlaIleHisTyrGluValHisProArgProGlyGlnAspGly 98	81 galAlaIleAlaIleHisTyrGluValHisProArgProGlyGlnAspGly 98
201 AGCTAATTCAGCCCATATATGAGGTTCATCTCGGCCAGACAGAGATGGAG 250	201 AGCTAATTCAGCCCATATATGAGGTTCATCTCGGCCAGACAGAGATGGAG 250
98 IeAlaIleAlaGlyValAspGlyThrValSerGlyTyrGluGluThrLysIle 114	98 IeAlaIleAlaGlyValAspGlyThrValSerGlyTyrGluGluThrLysIle 114
251 CACAAGCAGAGTGTGGATGGGACAGTGAATGCTGGGAAGAGACCAAAATC 300	251 CACAAGCAGAGTGTGGATGGGACAGTGAATGCTGGGAAGAGACCAAAATC 300
115 AspSerSerSerProLeuArgTyrAspArgGlnIleGlyGluPheThrVal 131	115 AspSerSerSerProLeuArgTyrAspArgGlnIleGlyGluPheThrVal 131
301 AACAGCTCCAGCCCTTCGCGCTACGACAGCCGACGATTTGGGAATTTACAGT 350	301 AACAGCTCCAGCCCTTCGCGCTACGACAGCCGACGATTTGGGAATTTACAGT 350
131 ILeArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluG 148	131 ILeArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluG 148
351 CATCAGAGCGTGGGCTACTACTGCTGACTGTCAGGTGACCTTTGATGAGG 400	351 CATCAGAGCGTGGGCTACTACTGCTGACTGTCAGGTGACCTTTGATGAGG 400
148 IyLysAlaValIlyrLeuLysLeuAspLeuValAsnGlyValLeuAla 164	148 IyLysAlaValIlyrLeuLysLeuAspLeuValAsnGlyValLeuAla 164
401 GAAAGGTGTCTACTAAGCTGAGACTGTGGTAAAGGTGTGGCGGCC 450	401 GAAAGGTGTCTACTAAGCTGAGACTGTGGTAAAGGTGTGGCGGCC 450
165 LeuArgCysLeuGluGluPheSerAlaThrAlaIleAspSerProGlyPro 181	165 LeuArgCysLeuGluGluPheSerAlaThrAlaIleAspSerProGlyPro 181
451 CTGCGCTCGCTGGAAGAATTTCTACGCCACAGCAGCAAGCTCTCTGGGCC 500	451 CTGCGCTCGCTGGAAGAATTTCTACGCCACAGCAGCAAGCTCTCTGGGCC 500
181 oGlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlyLys 198	181 oGlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlyLys 198
501 CGAGCTCCGTTTGTGGCAGGTGTGGGTGGCTGTGGCGCTGCGGCGAGGGGT 550	501 CGAGCTCCGTTTGTGGCAGGTGTGGGTGGCTGTGGCGCTGCGGCGAGGGGT 550
198 eSerLeuArgGlyLeuArgThrLeuProTyrAlaHisLeuLysAlaAlaPro 214	198 eSerLeuArgGlyLeuArgThrLeuProTyrAlaHisLeuLysAlaAlaPro 214
551 CTTCCCTCGGATTCGCAACCTCCCTCCGCGCATCTTAAGGCTGCCCCC 600	551 CTTCCCTCGGATTCGCAACCTCCCTCCGCGCATCTTAAGGCTGCCCCC 600
215 PheLeuThrTyrPheGlyLeuPheGlnValHis 225	215 PheLeuThrTyrPheGlyLeuPheGlnValHis 225
601 TTCCTAACCTACTTTGGACCTCTTCAATTCTCAC 633	601 TTCCTAACCTACTTTGGACCTCTTCAATTCTCAC 633
seq_name: /SID1/gcseqdata/geneseq/geneseqn-emb1/NA1998.DAT.AA47613	seq_name: /SID1/gcseqdata/geneseq/geneseqn-emb1/NA1998.DAT.AA47613
seq_documentation_block:	seq_documentation_block:
ID AA47613 standard: cDNA: 1236 BP.	ID AA47613 standard: cDNA: 1236 BP.
AC AA47613;	AC AA47613;
XX 27-OCT-1998 (first entry)	XX 27-OCT-1998 (first entry)
XX TNF related endothelium proliferative agent gene.	XX TNF related endothelium proliferative agent gene.
KW ss: TNF; endothelium proliferative agent; TREPA; wound healing	KW ss: TNF; endothelium proliferative agent; TREPA; wound healing
KM tissue grafting; vascularisation; apoptosis; autoimmune; diarr	KM tissue grafting; vascularisation; apoptosis; autoimmune; diarr
XX Homo sapiens.	XX Homo sapiens.
OS	OS
XX	XX
FH Key	FH Key
FT Location/Qualifiers	FT Location/Qualifiers
CDS 1..750	CDS 1..750
FT /tag= a	FT /tag= a
FT /product= "TREPA"	FT /product= "TREPA"
FT	FT

```

XX PN MO9835061-A2.
XX PD 13-AUG-1998.
XX PF 12-FEB-1998; 98WO-US02859.
XX PR 10-FEB-1998; 98US-0021706.
XX PR 12-FEB-1997; 97US-0798692.
XX PA (ABBO ) ABBOTT LAB.
XX PI
XX P1 WILEY SR.
XX DR WP1: 1998-447255/38.
XX DR P-PSDB; AAW29745.
XX PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
XX PT treatment of autoimmune disease, tumours and inflammation
XX Claim 11; Page 123-4; 142pp; English.
XX PS
XX PP The TNF-related endothelium proliferative agent (TREPA), or its
CC activators or agonists, are used to treat a deficit of TREPA, e.g. to
CC promote wound healing or tissue grafting, by promoting vascularisation,
CC also to induce apoptosis for treating cancer and eliminating autoreactive
CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
CC TRIPA peptides can also be used to target cytotoxic agents or for
CC affinity isolation of the corresponding receptor, the nucleic acid for
CC which can be used to transform tumour cells to render them more
CC responsive to TREPA and to screen for TREPA mimics.
CC Ribozymes, antisense RNA , antibodies or peptides, are used to treat
CC TREPA-associated inflammations, e.g. tumours and metastases (by inhibiting
CC vascularisation), inflammation or a wide range of autoimmune conditions,
CC conditions involving abnormal stimulation of epithelial cells (e.g.
CC atherosclerosis), for birth control (inhibiting ovulation and placental
XX formation) or other angiogenic conditions (e.g. ulcers).
XX SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

Alignment_scores:
Quality: 1020.00 Length: 224
Ratio: 4.744 Gaps: 0
Percent Similarity: 95.982 Percent Identity: 88.839

alignment_block:
US-09-245-198A-2 x AAV47613 ..

Align seg 1/1 to: AAV47613 from: 1 to: 1236

2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVa 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
76 CTGGGCCCTGGGCCTGGCCCTGGCCCTGGCCCTGGCCCTGGCGT 125
|||SerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
126 CAGTTTGGAACCGCGGCATCGTCGTCCGCCCAAGACCTGACCAGG 175
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
35 IuLeuThrAlaGluAspArgGluProPogIuLeuAsnProGlnThr 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
176 ACCTGGTGGCAGAGGAGCACGAGCCCGTGGAAATGCCACACA 225
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 GluGluSerGlnAspValValProPhelengLugInleuValArgProAr 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
226 GAAGAAGACCGAGATCTGGCCCTTTCTTCCAACCGACTAGTTGGCCTCG 275
gArgSerAlaProLySGlyArgLySalaArgProArgArgAlaIleaIaa 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
276 AAGAAGTGCACCTAAAGCCGGAACAACGCGGCTCGAAGAGCGATGCCAG 325
lanHtTyGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```
326 CCCATTATGAAAGTTTCACGACGACCTGGACAGGAGGAGCGGACGAGT 375
102 VALASPGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerse 118
376 GTGGAGGAGGACAGTAGTGCTGGGAGGAAGCCAGATCACAGCTCCAG 425
118 rProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaG 135
426 CCTCTGCGCTACACACCGCCAGATCGGGGAGTTTATATAGTCACCCGGGCTG 475
135 lYleuTyrTrpLeuTyrCysGlnValHisPheAspGluGlyLysAlaVal 151
476 GGCTCTACTACTGTACTGTACAGTGCACCTTGTATGAGGAGGAGGCTGTC 525
152 TyrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArgCysLe 168
526 TACCTAAGCTGAGCTTGTCTGGTGGATGGTGTCTGCTGCGCTGGCTGCC 575
168 uGluGluPheSerAlaThrAlaAlaSerSerProGlyProGluLeuArgL 185
576 GGAGGAATTCTCAGCCACTGGCGGAGATTCCCTCGGGCCCGCCAGCTCGCC 625
185 eucyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
626 TCTGCCAGGTGTCTGGCTGTGGCCCTCGCGGAGGAGGCTCTCCCTGCGG 675
202 lIeArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTy 218
676 ATCCGCAACCTCCCTCGGGCCATCTCAAGGCTGCCCGCTCTCTCAACCTA 725
218 rPheGlyLeuPheGlnValHis 225
726 CTTCGGAGCTCTTCACAGTTTCAC 747

seq_name: /SID1/gcdata/geneseq/geneseqn-emb1/NA2001A.DAT.AAD04350
seq_documentation_block:
ID AAD04350 standard; cDNA; 1236 BP.
AC AAD04350;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human TREPA (TNF related endothelium proliferative agent) cDNA.
XX
KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;
TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
grafting; vulnerrary; ss.
XX
SS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..750
FT FT /*tag= a
FT FT /product= "Human TREPA (TNF related endothelium
FT FT proliferative agent)"
XX
XX US6207642-B1.
XX
XX 27-MAR-2001.
XX
XX 26-JUN-1998; 98US-0105343.
XX
XX 12-FEB-1997; 97US-0798692.
XX
XX 10-FEB-1998; 98US-0021706.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Wiley SR;
XX
XX MPI: 2001-280760/29.
XX
XX P-PSDB; AAB00891.
XX
```

```
PT Inducing angiogenesis in mammal at desired sites for promoting wound
PT healing, by administering soluble fragment of extracellular domain of
PT tumor necrosis factor related endothelium proliferative agent protein
PT
PS Example 2: Column 73-74; 53pp; English.
XX
XX The present invention relates to extracellular signal molecules,
XX particularly members of tumour necrosis factor (TNF) family molecules
XX designated as TREPA (TNF related endothelium proliferative agent).
XX CC Soluble biologically active TREPA are used to treat TREPA-associated
XX diseases, tumours or metastases. TREPA is used for inducing angiogenesis
XX in human for promoting wound healing and for vascularizing grafted
XX CC tissue for successful grafting and to promote tissue grafts.
XX The present sequence is a cDNA clone ID #630050 encoding human TREPA.
SQ
Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:
```

```
alignment_scores:
Quality: 1020.00 Length: 224
Ratio: 4.744 Gaps: 0
Percent Similarity: 95.982 Percent Identity: 88.839
```

```
alignment_block:
US-09-245-198a-2 x AAD04350 ..
```

```
Align seg 1/1 to: AAD04350 from: 1 to: 1236
```

```
2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValVal 18
||||:|||||
76 CTGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 125
18 lSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
||||:|||||
126 CAGTTTGGGAGCGGGGATCGCTGTCGCGCCAGAGACCTGCCAGAGAG 175
35 luleuThrAlaGluAspArgArgGluProProGluLeuAsnProGlnThr 51
||||:|||||
176 AGCTGTGCGCAGAGAGAGACAGACCGCTCGGAGACTGAATCCCGAGACA 225
52 GluGluSerGlnAspValAlaProPheLeuGluGlnLeuValArgProAr 68
||||:|||||
226 GAAGAAAGCCAGAGATCCTGCGCTTCCTGTAACCGACTAGTTGGGCTCG 275
68 gArgSerAlaProLysGlyArgLysAlaArgProArgArgAlaIleAla 85
||||:|||||
276 AAGAAAGTGCACCTAAAGCGCGGAAACACGCGGCTCGAAGAGCATGCGAC 325
85 lAhisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
||||:|||||
326 CCCATTATGAAAGTTTCACGACGACCTGGACAGGAGGAGCGAGCGAGGT 375
102 VALASPGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerse 118
||||:|||||
376 GTGGAGGAGGACAGTAGTGCTGGGAGGAAGCCAGATCAACAGCTCCAG 425
118 rProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaG 135
||||:|||||
426 CCTCTGCGCTACACACCGCCAGATCGGGGAGTTTATATAGTCACCCGGGCTG 475
135 lYleuTyrTrpLeuTyrCysGlnValHisPheAspGluGlyLysAlaVal 151
||||:|||||
476 GGCTCTACTACTGTACTGTACAGTGCACCTTGTATGAGGAGGAGGCTGTC 525
152 TyrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArgCysLe 168
||||:|||||
526 TACCTGAAGCTGAGCTTGTGTGGATGGTGTCTGCTGCGCTGGCTGCC 575
168 uGluGluPheSerAlaThrAlaAlaSerSerProGlyProGluLeuArgL 185
||||:|||||
576 GGAGGAATTCTCAGCCACTGCGGCGAGTTCTCTCGGGCCCGCCAGCTCGCC 625
```

185 euYsglValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
|||||
626 TCTGCAGGTGCTGGGCTGTGGCCCTGCGCCAGGCTCTCCCTGGCG 675
202 lIeaGThrLeuProTIPAlaHisLeuLysAlaAlaProPheLeuThy 218
|||||
676 ATCCGACCCCTCCCTGGCCCATCTCAAGGCTGCCCTTCCTCACTTA 725
218 rPheGlyLeuPheGlnValHis 225
|||||
726 CTTCGACTCTTCAGGTTCAAC 747

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA49717

seq_documentation_block:

ID AAA49717 standard; cDNA; 1353 BP.
XX
AC AAA49717;
XX
DT 25-SEP-2000 (first entry)
XX
DE Human PRO207 cDNA clone DNA30879-1152.
XX
KW PRO207; human; antitumour; tumour; therapy; cytostatic;
KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;
KW uterine cancer; prostate cancer; lung cancer; bladder cancer;
KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 58..807
FT /*tag- a
FT sig_peptide 58..177
FT /*tag- b
FT mat_peptide 178..804
FT /*tag- c
XX
PN WO200037638-A2.
XX
PD 29-JUN-2000.
XX
PE 02-DEC-1999; 99MO-US28565.
XX
PR 22-DEC-1998; 98US-0113296.
PR 08-MAR-1999; 99MO-US05028.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 15-SEP-1999; 99MO-US21090.
PR 15-SEP-1999; 99MO-US21547.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;
PI Napier MA, Picti RM, Wood WI;
XX
DR WPI; 2000-442668/38.
DR P-SDS; AAY95338.
XX
XX Novel composition to inhibit neoplastic cell growth or for treating
PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO509 or
PT PRO866
XX
PS Claim 20; Fig 3; 172pp; English.
XX
CC The present sequence is that of cDNA clone DNA30879-1152
CC (AYCC 209358) encoding human PRO207 (see AAY95338), which shows
CC homology to several members of the tumour necrosis factor family,
CC especially human lymphotoxin (23.4%). The cDNA was identified in a

CC foetal kidney cDNA library following identification of an expressed
CC sequence tag with homology to human Apo-2 ligand. A claimed method
CC for inhibiting the growth of a tumour cell comprises exposing the
CC tumor cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,
CC PRO328, PRO301, PRO526, PRO362, PRO509 or PRO866 (see
CC AAY95337-49), their agonists or chimeric polypeptides incorporating
CC them. The tumour is especially a cancer selected from breast,
CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and
CC central nervous system cancer, melanoma and leukemia. Nucleic
CC acids encoding PRO179 etc. are used in the recombinant production
CC of the antitumour polypeptides.
XX
XX Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

alignment_scores:

Quality: 1020.00 Length: 224
Ratio: 4.744 Gaps: 0
Percent Similarity: 95.982 Percent Identity: 88.839

alignment_block:

US-09-245-198a-2 x AAA49717 ..

Align seg 1/1 to: AAA49717 from: 1 to: 1353

2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVal 18
|||||
133 CTGGGCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 182
18 lSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||
183 CAGTTGGGGAGCCGCGCATCGCTCCGCCAGAGCGTCCAGAGAG 232
35 lLeuThrAlaGluAspArgArgGluProGluLeuAsnProGlnThr 51
|||||
233 AGCTGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 282
52 GluGluSerGlnAspValValProPheLeuGluGlnLeuValArgPro 68
|||||
283 GAG 332
68 ArgSerAlaProLysGlyArgGlyAlaArgProArgAlaAlaAla 85
|||||
333 CAGAGTGACACTAAAGCCGGAACAAACAGGCTCGAAGAGAGATCGCG 382
85 lAlaIleTyrGluValHisProArgProGlyGlnAspGlyAlaGlnA 101
|||||
383 CCCATTATGAAGTTCATCCAGACCTCGAGAGAGAGAGAGAGAGAGAG 432
102 ValAspGlyThrValSerGlyTyrPheGluGluThrLysIleAsnSer 118
|||||
433 GTGAG 482
118 rProLeuArgTyrAspArgIleIleGlyGluPheThrValIleArgAla 135
|||||
483 CCTCTGGCGCTACACACCGCAGATGGGGAGTTTATGCAACCGCGCTG 532
135 lLeuLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyAlaVal 151
|||||
533 GGCTCTACTACTCTGACTGTCAGTGCACCTTGTGAGAGAGAGAGAGCT 582
152 TyrLeuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCys 168
|||||
583 TACCTGAAGCTGACTCTGTGTGATGGATGGTGGCTGGCTGGCTGGCT 632
168 uGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArg 185
|||||
633 GGAGGAATTCACGCACTCGCGGAGATTCCCTCGGCGCCAGATCCGCC 682
185 euYsglValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
|||||
683 TCTGCAGGTGCTGGCTGTGGCCCTGCGGCCAGGCTCTCCCTGGCG 732

202 ILeaRgThLeuProTrrPaLaHIsLeuLySaLaAaProPhLeuThrTy 218
|||||
733 ATCCGACCTCCCTCCGCGCCATCTCAAGGCTGCCCTCTCTCACCTA 782
218 rPhegLyLeuPhegInValHis 225
|||||
783 CTTCGGACTCTTCCAGGTTTAC 804

seq_name: /SIDS1/gcdata/geneseq/geneseqn-emb1/NA1998.DAT.AAV18600

seq_documentation_block:

ID AAV18600 standard; CDNA: 1373 BP.

AAV18600;

21-JUL-1998 (first entry)

XX Homo sapiens tumour necrosis factor related ligand (TRELL) gene.

XX TRELL; tumour necrosis factor related ligand; tnfr; treatment;
cancer; autoimmune disease; immune system; stimulation; suppression;

XX graft rejection; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..852

FT /tag= a

FT /note= "tumour necrosis factor related ligand"

XX MO9805783-A1.

XX 12-FEB-1998.

XX 07-AUG-1997; 97MO-US13945.

XX 18-MAR-1997; 97US-0040820.

XX 07-AUG-1996; 96US-0023541.

XX 18-OCT-1996; 96US-0028515.

XX (BIOJ) BIOGEN INC.

XX (UYGE-) UNIV GENEVA FACULTY MEDICINE.

XX Browning JL, Chicheportiche Y;

XX WPI: 1998-145619/13.

XX P-PSDB: AAW47525.

XX Tumour necrosis factor related ligand - useful for, e.g. treating
cancer, auto-immune disease and immune responses to tissue grafts

XX Claim 2; Pages 48-50; 69pp: English.

XX The sequence is that encoding human tumour necrosis factor related
ligand (TRELL). TRELL or active fragments can be included with a
carrier in pharmaceutical compositions to treat cancer, autoimmune
diseases or immune responses to tissue grafts, or to stimulate or
suppress the immune system. It is useful to screen for TRELL
receptors, by labelling with a detectable label and screening
compositions for binding. Agents interfering with TRELL-receptor
binding can also be screened for, can then be administered,
optionally with interferon- gamma, to induce cell death or
treat, suppress or alter immune responses (especially involving human
adenocarcinoma cells) involving a signal pathway between TRELL and its
receptor. The DNA sequence can be used in gene therapy for

XX TRELL-related disorders in mammals (especially humans), e.g. tumours,
autoimmune and inflammatory diseases or inherited genetic disorders,
by introducing into cells, and expressing, therapeutically effective
amounts of a vector, e.g. a virus comprising a gene encoding TRELL.

XX It may also be of use in the preparation of gene encoding probes for
screening natural/synthetic DNAs for TRELL-encoding sequences
and for antisense therapy.

XX seq_name: /SIDS1/gcdata/geneseq/geneseqn-emb1/NA1999.DAT.AAX56000

seq_documentation_block:

ID AAX56000 standard; DNA: 1421 BP.

XX

S0 Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other:

alignment_scores:

Quality: 1020.00

Ratio: 4.744

Percent Similarity: 95.982

Percent Identity: 88.839

alignment_block:

US-09-245-198a-2 x AAV18600

Align seg 1/1 to: AAV18600 from: 1 to: 1373

2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVala 18
|||||
181 CTGGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCT 230
18 IserLeuGlySerTrrPaLaHIsLeuSerAlaGlnGluProSerGlnGlu 35
|||||
231 CAGTTTGGGGAGCCGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 280
35 ILeuThrAlaGlnAspArgArgGluProGluLeuAsnProGlnThr 51
|||||
281 AGCTGTGGCAGAGAGAGACAGACAGACCGCTCGAATGATCCACAGACA 330
52 GluGluSerGlnAspValValProPhLeuGlnGlnLeuValArgProAr 68
|||||
331 GAAGAAAGCCAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380
68 GarGSerAlaProLyGlyArgLySaLaArgProArGArgAlaIleAlaA 85
|||||
381 CAGAACTGCACCTAAAGCCGGAACACAGGCGCTCCAGACGAGCTGCGAG 430
85 IahIstYrGluValIHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
|||||
431 CCCATTATGAAAGTTTCATCCACGACCTGGACAGGAGCGACGACGAGGT 480
102 ValAspGlyThrValSerGlyTrrPGLuGlnThrLysIleAsnSerSerSe 118
|||||
481 GTGGAGCGGACAGTGAAGTGGCTGGAGGAGGAGGAGGAGGAGGAGGAG 530
118 rProLeuArgTyrAspArgGlnIleGlyGluPhMetThrValIleArgAlaG 135
|||||
531 CCTCTGCGCTTACACCGCCAGATCGGGAGTTTATATAGTCACCCGGCTG 580
135 ILeuThrTyrLeuTyrCysGlnValIHisPheAspGlyGlySaLaVala 151
|||||
581 GGCTTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630
152 TyrLeuLysLeuAspLeuValaGlnGlyValLeuAlaLeuArgCysLe 168
|||||
631 TACCTGAAGCTGACCTGGTGGATGGTGGTGGTGGTGGTGGTGGTGGCT 680
168 uGluGluPheSerAlaThrAlaIleSerSerProGlyProGlnLeuArgL 185
|||||
681 GGGGGAATTTCTGAGCCAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 730
185 euCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
|||||
731 TCTGCCAGGTGTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 780
202 ILeaRgThLeuProTrrPaLaHIsLeuLySaLaAaProPhLeuThrTy 218
|||||
781 ATCCGACCTCCCTCCGCGCCATCTCAAGGCTGCCCTCTCTCACCTA 830
218 rPhegLyLeuPhegInValHis 225
|||||
831 CTTCGGACTCTTCCAGGTTTAC 852

XX AAX56000;
 AC
 XX
 DT 15-JUL-1999 (first entry)
 XX
 DE Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.
 XX
 KW Human: tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
 KW NP-kappaB-dependent transcription; JNK/SAPK-dependent response;
 KW cancer; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT CDS 92..841
 FT /*tag= a
 FT /product= "Apo-3 ligand"
 PN MO9919490-A1.
 XX
 D 22-APR-1999.
 XX
 PE 09-OCT-1998; 98WO-US21407.
 XX
 PR 17-DEC-1997; 97US-0069862.
 PR 10-OCT-1997; 97US-0062037.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Marsters SA, Pitti R;
 XX
 DR WPI; 1999-287982/24.
 DR P-PSDB; AAY09369.
 XX
 PT New human Apo-3 ligand (a tumour necrosis factor) homologue
 PS
 PS Claim 18; Fig 1; 74pp; English.
 XX
 CC The present sequence encodes a human tumour necrosis factor (TNF) and
 CC lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
 CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis in
 CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and
 CC to induce JNK/SAPK-dependent responses in mammalian cells.
 XX
 SO Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other;

alignment_scores:

Quality: 1020.00 Length: 224
 Ratio: 4.744 Gaps: 0
 Percent Similarity: 95.982 Percent Identity: 88.839

alignment_block:

US-09-245-198a-2 x AAX56000 ..

Align seg 1/1 to: AAX56000 from: 1 to: 1421

2 leuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVala 18
 |||
 167 CTGGCCCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 216
 18 lserLeuGlySerTPAlaThrLeuSerAlaGlnGlnProSerGlnGln 35
 |||
 217 CAGTTGGGGACCGCGCATCGCTGCCCGAGGAGCGCTGGCAGAGG 266
 35 luleThrAlaGlnAspArgArgGlnProGlnLeuAsnProGlnThr 51
 |||
 267 AGCTGTGGCAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 316
 52 GluGlnSerGlnAspValValProPheLeuGlnGlnLeuValAlaArgProAr 68
 |||
 317 GAAGAAAGCCAGAGATCTGGCCCTTCTGAAACGACTAGTTGGGCTCG 366

68 garSerAlaProLysGlyArgLysAlaArgProArgAlaIleAla 85
 |||
 367 CAGAGTGACCTAAAGCGGAAACAGCGGCTCGAAGAGGATCGCAG 416
 85 laHisTyrGlnValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
 |||
 417 CCCATTATGAAAGTCAATCCACGACCTGCGACAGCAGCAGCAGCAGCAG 466
 102 ValAspGlyThrValSerGlyTyrPgluGlnThrLysIleAsnSerSerSe 118
 |||
 467 GTGAGCGGACAGTAGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 516
 118 rProLeuArgTyrAspArgGlnIleGlyGlnPheThrValIleArgAlaG 135
 |||
 517 CCTCTGGCGCTACAAACCGCAGATCGGGAGTTTATAGTACCCGGGCTG 566
 135 lLeuTyrTyrLeuTyrCysGlnValHisPheAspGlnGlyLysAlaVal 151
 |||
 567 GCGCTACTACTGCTGACTGTGCTGATGCTGCTGCTGCTGCTGCTGCTG 616
 152 TyrLeuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLe 168
 |||
 617 TACCTGAAGCTGACTTGTGCTGGATGCTGCTGCTGCTGCTGCTGCTG 666
 168 uGlnGlnPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgL 185
 |||
 667 GGAGGAATTTCTAGCCACTGCGGCGGAGTTCCCTCGGCGCCCGCTCCGC 716
 185 euCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
 |||
 717 TCTGCCAGGTGCTGGGCTGTGGCCCTGGCCAGGAGGCTCCCTCGCGG 766
 202 lIleArgThrLeuProThrAlaHisLeuLysAlaAlaIleProPheLeuThrTy 218
 |||
 767 ATCCGACCCCTCCCTGGGCGCATCTCAAGGCTGCCCTTCCTCACTCA 816
 218 rPheGlyLeuPheGlnValHis 225
 |||
 817 CTTGGACTCTTCCAGGTTTCAC 838
 seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: AAS03964
 seq_documentation_block:
 ID AAS03964 standard; DNA; 898 BP.
 AC AAS03964;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 XX Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.
 DE
 XX
 XX TWEAK extracellular domain: tumour necrosis factor; TNF; angiogenesis;
 KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
 KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
 KW rubrosis; uveitis; macular degeneration; arthritis; rheumatism; ds;
 KW corneal graft neovascularisation; psoriasis; metastatic condition;
 KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
 KW preneoplastic condition; myocardial angiogenesis; wound granulation;
 KW scleroderma; vascular adhesion; telangiectasia; ischemia; human;
 KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;
 KW peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;
 KW fusion protein.
 XX
 XX Homo sapiens.
 OS
 OS Synthetic.
 XX
 FH Key location/Qualifiers
 FT CDS 52..873
 FT /*tag= a
 FT /product= "Fusion protein comprising a growth hormone
 leader, a leucine zipper multimerisation
 domain, and human TWEAK extracellular
 domain"

XX WO200145730-A2.
XX 28-JUN-2001.
XX
XX 19-DEC-2000; 2000WO-US34755.
XX
XX 20-DEC-1999; 99US-0172878.
XX 10-MAY-2000; 2000US-0203347.
XX
XX (IMV) IMMUNEX CORP.
XX
XX Wiley SR;
XX WPI; 2001-417975/44.
XX P-PSDB; AAU03499.
XX
XX Modulating angiogenesis in a mammal for treating diseases mediated by
XX angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
XX peripheral tissue, by administering antagonist or agonist of TWEAK
XX receptor

PS Example 1: Page 39-40; 46pp; English.

XX The sequence represents a DNA from the expression vector
XX PDE409-LZ-TWEAK, which encodes a fusion protein comprising a growth
XX hormone leader, a leucine zipper multimerisation domain, and the
XX extracellular domain of human TWEAK. The fusion protein was used in
XX the isolation of human TWEAK receptor (TWEAKR)-expressing clones
XX from a COS cell human cDNA library. The TWEAK protein is a
XX member of the tumour necrosis factor (TNF) family and induces
XX angiogenesis. TWEAKR may therefore be used to screen for and develop
XX TWEAKR agonists and antagonists for the modulation of angiogenesis, to be
XX used in the treatment and diagnosis of human disease. The disorders
XX mediated by angiogenesis include ocular disorders characterised by ocular
XX neovascularisation such as diabetic retinopathy, neovascular glaucoma,
XX retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
XX rubeosis, uveitis, macular degeneration and corneal graft
XX neovascularisation, and inflammatory diseases such as arthritis,
XX rheumatism and psoriasis. Other treatable diseases include malignant and
XX metastatic conditions such as sarcomas and carcinomas, benign tumours and
XX preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
XX scleroderma, vascular adhesions, atherosclerotic plaque
XX neovascularisation, telangiectasia, wound granulation, coronary
XX atherosclerosis, peripheral atherosclerosis and ischaemia.

XX Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other;

XX Alignment_scores:

XX Quality: 951.00 Length: 207
XX Ratio: 4.803 Gaps: 0
XX Percent Similarity: 95.652 Percent Identity: 88.889

XX Alignment_block:

XX US-09-245-198a-2 x AAS03964 ..

XX Align seg 1/1 to: AAS03964 from: 1 to: 898

19 SerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||
250 AGTTTGGGGAGCCGGCAGTCGTCGCCAGAGAGCCGTCAGAGGGA 299
35 uLeuThrAlaGluAspArgGluProProGluLeuAsnProGlnThrG 52
|||||
300 GCGTGCGGCGAGAGAGACAGACCCGTCGGAACGAACTGACCCAGACAG 349
52 LuGluSerGlnAspValValProPheLeuGlnGluLeuValAlaArgProArg 68
|||||
350 AAGAAAGCCGAGATCTGCGCTTTCCTGAACGACGACGAGTTGCGCTCGC 399
69 ArgSerAlaProGlyGlyArgGlyAlaArgProArgArgAlaAlaAla 85
|||||

400 AGAAGTGACACTAAAGCCGAGAAACAGCGGCTCGAAGAGCATCGCACC 449
85 aHsTYrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 102
|||||
450 CCATTTATGAAGTTTCATCCAGACCTCGACAGAGAGCGAGCGAGGTTG 499
102 aLAspGlyThrValSerGlyTYrPGLuGluThrLysLLeasSerSerSer 118
|||||
500 TGGACGGAGACAGTGAAGTGGCTGGGAGAAAGCCAGAAATCAACAGCTCCAC 549
119 ProLeuArgTYrAspArgGlnIleGlyGluPheThrValIleArgAlaG 135
|||||
550 CCTTGGCGCTACAAACCCAGATCGGGGAGTTTATGTGACCCCGGGCTGG 599
135 yLeuTYrTYrLeuTYrCysGlnValHisPheAspGluGlyLysAlaValT 152
|||||
600 GCTCTACTACTGTACTGTACAGGTGCATTTGATGAGGGGAAGCGTGTCT 649
152 yTLeuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLeu 168
|||||
650 ACCTGAAGCTGGAGCTTGCTGTGATGTGTGCTGCTGCGCTGCTGCTG 699
169 GluGluPheSerAlaThrAlaAlaSerProGlyProGlnLeuArgLe 185
|||||
700 GAGGAATTCACAGCCACTGGGCGCAGTTCCCTCGGCGCCACCTCGCCT 749
185 uCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 202
|||||
750 CTCGCAAGTGTCTGGGCTGTGTGGCCCTCGCGGCAAGGCTCTCCCTCGGA 799
202 leArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTYr 218
|||||
800 TCCGCAACCTCCCTGGGCCATCTCAAGGCTGCCCTTCCTCAGACCTAC 849
219 PheGlyLeuPheGlnValHis 225
850 TTCGACTCTTCAGGTTTCAC 870

seq_name: /SIDSL/gcdata/geneseq/geneseqn-embL/NA1999.DAT:AAx23424

seq_documentation_block:

ID AAX23424 standard; DNA; 1030 BP.

XX AAX23424;

XX 18-JUN-1999 (first entry)

XX Human TNRL3 DNA.

XX Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;
XX developmental abnormality; gestational abnormality; prostate cancer;
XX AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
XX cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
XX apoptosis; human; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS

XX 1..627

XX /*tag= a

XX /product= "TNRL3"

XX WO9911791-A2.

XX 11-MAR-1999.

XX 04-SEP-1998; 98WO-US18393.

XX 05-SEP-1997; 97US-0924634.

XX (UNITW) UNITV WASHINGTON.

XX Chaudhary PM;

S0 Sequence 978 BP; 289 A; 261 C; 238 G; 190 T; 0 other:

alignment_scores:

Quality: 110.50 Length: 197
Ratio: 0.978 Gaps: 10
Percent Similarity: 57.360 Percent Identity: 24.873

alignment_block:
US-09-245-198a-2 x ABL21473 ..

Align seg 1/1 to: ABL21473 from: 1 to: 978

```

41 ArgArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspVa 57
   ::::::::::::::::::::|
450 CAAAGAGACCCCTGCACCACTTCACCA.....CCGTCGCAAGATGC 490
   |||
57 lvalPro.pheLeuGluGlnLeuValArgProArgArgSer.....Ala 71
   :|||
491 ATTCCCGCATCGCCACCTCTAGTCGCAAGAGTGAATCTCTTTTCA 540
   :|||
72 ProLysGlyArgLysAlaArgProArgArgAlaIleAlaIleHisTyrG1 88
   ::::::::::::::::::::|
541 GCCAGATCGAGAGCTCGAGGCCA.....GCAGCCCATTTCCA 578
   |||
88 uValHisProArgProGluGlnAspGlyAlaGlnAlaGlyValAspGlyT 105
   ::::::::::::::::::::|
579 CTTGACACACAGCGCGCTGACCAAGAGATAG...GGCTACCATGAGAG 625
   |||
105 hValSerGlyTyrGluGluThrLysIleAsnSerSerProLeuArg 121
   ::::::::::::::::::::|
626 ATATGTACATAGGAATGATTAACGAGAAACTCTTATCAGGA...CAC 672
   |||
122 TyrAspArgGlnIleGlyLuphethValIleArgAlaGlyLeuTyrTy 138
   ::::::::::::::::::::|
673 TTTCAAACCGCGATGCGCTTTGACGGTACCAATACAGAGCCTATATTA 722
   |||
138 rLeuTyrGysGlnValHisPheAspGluGlyLysAlaValTyrLeuLys1 155
   ::::::::::::::::::::|
723 CGATATCGCCAGATATGTCTACCAACATCGCAGAC..... 759
   |||
155 euAspLeuLeuValAsnGlyValLeuAla..... 164
   |||
760 .....CAGAACGATTTATGCTCTTCAAGAGACACTCCATTC 798
   |||
165 LeuATGcysLeuGluGluLupheserAlaThrAlaAlaSerSerProGlyPr 181
   |||
799 CTGCACTGCTTGAAAC.....ACGTGCCCAACCAACATCGGCACA 836
   |||
181 oGlnLeuATGcysGlnValSerGlyLeuLeuProLeuArgProGlyS 198
   ::::::::::::::::::::|
837 TAAAGTGCAACCTGCCACACGAGTGTCTGATCCACTGGAAAGAAACG 886
   |||
198 eSerLeuATGcysGlnValSerGlyLeuLeuProLeuArgProGlyS 211
   ::::::::::::::::::::|
887 AGAGATTCATCTGAAGACATTCACACAGATCGCAATGCATGCTTGGCG 936
   |||
212 AlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 244
   ::::::::::::::::::::|
937 GAGGGAACACCGAGCTACTTTGGCATCTTCAAGGTG 975
   |||
seq_name: /SIDSL/gcdata/geneseq/geneseq-emb1/NA1998.DAT.AAV41377
seq_documentation_block:
ID AAV41377 standard; cDNA: 1630 BP.
AC AAV41377;
XX
XX
XX 08-OCT-1998 (first entry)
XX
XX NF-kB receptor activator RANK ligand (RANKL) encoding cDNA.
XX
XX RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;
KW

```

```

KW immune response; inflammatory response; toxic shock; sepsis;
KW RANKL; RANK ligand; tumour necrosis factor; TNF; mouse; ss.
XX
XX Mus musculus.
OS
XX Key Location/Qualifiers
XX Key 3..887
XX CDS /*tag= a
XX FT /product= "murine RANKL (ligand for RANK)"
XX
XX WO9828426-A2.
XX
XX PD 02-JUL-1998.
XX
XX PF 22-DEC-1997; 97WO-US23775.
XX
XX PR 14-OCT-1997; 97US-0064671.
XX PR 23-DEC-1996; 96US-0059978.
XX PR 07-MAR-1997; 97US-0813509.
XX
XX (IMMV ) IMMUNEX CORP.
XX
XX PI Anderson DM, Galibert LJ, Maraskovsky E;
XX
XX WP1: 1998-377657/32.
XX
XX DR P-PSDB; AAW69956.
XX
XX PT New isolated ligand for receptor activator of NF-kappa B - used to
XX develop products for augmenting an immune response for inhibiting an
XX inflammatory response and for protection of cells
XX
XX Claim 25; Pages 55-57; 80pp; English.
XX
XX This cDNA encodes a murine RANKL, a ligand for the RANK (receptor
XX activator of necrosis factor-kappa B (NF-kB)) polypeptide. RANK is a
XX member of the tumour necrosis factor (TNF) family. A soluble RANK
XX may be used for inhibiting activation of NF-kB, by contacting a cell
XX expressing membrane-associated RANK with a soluble RANK which binds to
XX RANK ligand (RANKL). RANKL polypeptides can activate RANK and can be
XX used to induce maturation of dendritic cells and enhance their
XX CC soluble RANK polypeptide composition may also be used for regulating
XX the immune or inflammatory response. Inhibition of NF-kB by RANK antagonists
XX may be useful in ameliorating negative effects of an inflammatory
XX response that result from triggering of RANK, e.g. in treating toxic
XX shock or sepsis, graft-versus-host reactions, or acute inflammatory
XX reactions. They can also be used in adjunct therapy for disease
XX characterised by neoplastic cells that express RANK. RANKL polypeptides
XX can also be used to identify inhibitors of RANK and thus inhibitors of
XX an inflammatory response, and also for protecting RANK-expressing cells
XX from the negative effects of chemotherapy or the presence of high levels
XX of TNF-alpha. The products can also be used for detection and drug
XX screening.
XX
XX S0 Sequence 1630 BP; 436 A; 355 C; 379 G; 460 T; 0 other:

```

alignment_scores:

Quality: 106.50 Length: 276
Ratio: 0.772 Gaps: 13
Percent Similarity: 50.000 Percent Identity: 24.275

alignment_block:

US-09-245-198a-2 x AAV41377 ..

Align seg 1/1 to: AAV41377 from: 1 to: 1630

```

3 SerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValAlaSe 19
   |||
78 TCCATGTTCTGCTGCTCGGCGGAGCTGGGCAAGTGGCTGTCGAG 127
   |||
19 rLeuGlySerTriPalaThrLeuSerAlaGln...GluProSerGlnGluG 35
   |||

```

```

128 CATCGGCTCTGTCCTCTACTTTCAGGCGCAGATGGATCCTAACAGA.... 173
35 IuleuthralagluAspArg..... 41
174 ....ATATCAGAGACGAGCAGCTCACTGCTTTATAGAACCTCGAGACTC 218
42 ArgGluProProGluIuleuasnProGlnthrGluIuseGlnAspVala 58
219 CATGAAACCGACATTTTGCAAGACTCGACTGTGGAGAGTGAAGACACT 268
58 lPro.....59
269 ACCTGACTCCTCGAGAGATGAACAAGCCTTTCAGGGGGCCTGCAGA 318
60 ..PheIugIugInIeuValaIarProArg.....SerIaIarPoly 73
319 AGCACTGCAACACATTGTGGGGCCACAGCCCTTCTCAGGAGCTCAGCT 368
74 .....GlyArgIysAlaArgProArgAr 81
369 ATGATGGAAGCGTCATGTTGGATGTGGCCGACGAGCAACCGCTGAGGC 418
81 gAlaIleAlaIaHisItrGluValaHisPProArgProGlyGlnAspGlyA 98
419 CCAGCCATTGGACACCTCCACCATCAACAAAGTCGCCAGCATCCCATGGGTT 468
98 lagIlnaIagIyAlaSpGIyThraIserGIyTrpGlu..... 111
469 CCAT.....AAGTCACCTGTCTCTGTGTACACAGCATGAGGC 509
112 ...ThIyIleAsnSerSerSerProIeuArgIyAspArgIlnIleGI 127
510 TGGGCGCAAGATCTCTAACATGACG.....TTAAGCAACGG 544
127 yGIuPheThValIleArgAlaGlyLeuTyTrIeuTy.....C 141
545 AAACTAAGAGTTAAACCAAGATGGCTTCTATTACCTGTACGCCAACATTT 594
141 yGlnIValaHisPheAspGluGlyIysAlaVal.....TyIleuIys 154
595 GCTTGGCGCATCATGAACAATCGGGAAGCGTACCTACAGACTATCTTCAG 644
155 IeuAspIeuVal.....AsnGlyValIleuAlaIeuArgCy 167
645 CTGATGCTGTATGTCGTTAAACCCAGCATCAAAATCCCAAGTTTCATGA 694
167 sleuGIuGluPheSerAlaThrAlaIaIaSerSerProGIyProGlnIlea 184
695 CCGATGAAAGAGGAGGACAGAAAACCTGGTCGGGCATTTCTGATTC 744
184 rglu.....CysGlnValaSerGlyLeuIeuProIeuArgProGIySer 198
745 ACTTTATTCATGAAGAATGTGGGAGATTTTTCAGAGCTCGAGCTGTGAA 794
199 SerIleaArgIleArgThrIleuProTrpAlaHisIleuIysAlaIaIaIaProh 215
795 GAAATTAAGCATTCAGGTGTCCAAACCTTCCTGCTGATCCGATCAAGA 844
215 eleuthrTyrrPheGlyIeuPheGlnVal 224
845 TCGGAGCTACTTGTGGGCTTCAAAAGTT 872

seq_name: /SIDS1/gcgcdata/geneseq/geneseqn-emb1/MA1998.DAT.AA413711
seq_documentation_block:
ID AA41371 standard; CDNA: 1630 BP.
XX AA41371:
XX AC
XX AA41371:
XX DT 08-Oct-1998 (first entry)
XX DE NF-kB receptor activator RANK ligand (RANKL) encoding cDNA.

```

XX	MMNK: necrosis factor-kappa B; NF-kB: receptor activator; human;
KM	immune response; inflammatory response; toxic shock; sepsis;
KW	RANKL; RANK ligand; tumour necrosis factor; TNF; mouse; ss.
XX	
OS	Mus musculus.
XX	
FH	Key
FT	Location/Qualifiers
CDS	3..887
FT	/*tag= a
FT	/product= "murine RANKL (ligand for RANK)"
XX	
PN	WO9828424-A2.
XX	
PD	02-JUL-1998.
XX	
PE	22-DEC-1997; 97WO-US23866.
XX	
PR	14-OCT-1997; 97US-0064671.
PR	23-DEC-1996; 96US-0059978.
PR	07-MAR-1997; 97US-0813509.
XX	
PA	(IMMV) IMMUNEX CORP.
XX	
PI	Anderson DM, Galibert LJ, Maraskovsky E;
XX	
DX	WPI; 1998-377655/32.
DR	P-PsDB; AAW68292.
XX	
PT	New isolated receptor activator of necrosis factor-kappa B - useful
PT	for, e.g. developing products for regulating an immune or
PT	inflammatory response, treating toxic shock or sepsis
XX	
DS	Example 7; Pages 55-57; 80pp; English.

CC This cDNA encodes a murine RANKL, a ligand for the RANK (receptor
CC activator of necrosis factor- κ B) polypeptide. RANK is a
CC member of the tumor necrosis factor (TNF) family. Host cells transformed
CC or transfected with an expression vector comprising the RANK encoding
CC nucleic acid can be used to produce recombinant RANK protein. The soluble
CC RANK may be used for inhibiting activation of NF- κ B, by contacting a cell
CC expressing membrane-associated RANK with a soluble RANK which binds to
CC RANK ligand (RANKL). The soluble RANK polypeptide composition may also be
CC used for regulating an immune or inflammatory response. Inhibition of
CC NF- κ B by RANK antagonists may be useful in ameliorating negative effects
CC of an inflammatory response that result from triggering of RANK, e.g. in
CC treating toxic shock or sepsis, graft-versus-host reactions, or acute
CC inflammatory reactions. They can also be used in adjunct therapy for
CC disease characterised by neoplastic cells that express RANK. The products
CC can also be used for detection and drug screening.

SQ Sequence 1630 BP; 436 A; 355 C; 379 G; 460 T; 0 other;

XX

Alignment_scores:

	Quality:	106.50	Length:	276
	Ratio:	0.772	Gaps:	13
	Percent Similarity:	50.000	Percent Identity:	24.275

alignment_block:
US-09-245-198A-2 x AAIV41371 ..

Align seg 1/1 to: AAIV41371 from: 1 to: 1630

```
3 SerleuglyleuAlaleuAlaCysleuglyleuleuValvalalse 19
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
78 TCcATGTTCCCTGGGCCCTCCCTGGGGCTGGAGTCAGCGCAGGTGCTGCACg 127
19 rleuglyserTrpPalaThrleuseSerAlagin...Gluproserglnglug 35
|:::~::~ :::: ||||| ::|||:::~:::
128 CATGCGTCTGTCTCGACTTTCAGAGCGCAGATGATCCTAACAga.... 173
35 luleurhralagluasparg..... 41
```

```

174 ....ATATCAGAGACAGCAGCTCACTGCTTTTATAGANTCTGAGACTC 218
XX
42 ATGGJUPProgluLeuasnProgluThrgluSergluAspVala 58
XX
219 CATGAAACGACAGATTTCGAGACTGCTGAGAGTGAAGACACT 268
XX
58 LPro..... 59
XX
269 ACCTGACTCTGCGAGAGATGAACACGCTTTCAGGGGCGGTGACA 318
XX
319 AGCACTGCAACACATTTGGGGCCACAGCCTTCTCAGAGACTCCAGCT 368
XX
74 .....GlyArgLysAlaArgProArg 81
XX
369 ATGATGGAAGCTCATGTTGGATGTGGCCGACGAGGACAGCTGAGGC 418
XX
81 gAlaIleAlaIleHisTyrGluValHisProArgProgluAspGly 98
XX
419 CCAGCATTTGGCACACCTCACCATCAATGCTGCAGCATCCCATCGGTT 468
XX
98 IagIleAlaIleValAspGlyThrValSerglyTyrGluVal..... 111
XX
469 CCCAT.....AAAGTCACTCTGCTCTTGTACACGACATCGAGGC 509
XX
112 ...ThrLysIleasnSerSerProleuArgTyrAspArgIleIle 127
XX
510 TGGGCCAAGATCTCTACATGACG.....TTAAGCACGCG 544
XX
127 yGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr..... 141
XX
545 AAACACTAAGGGTTAAACAGATGGCTTTATACCTGACGCCAACATTT 594
XX
141 yGluValHisPheAspGluGlyLysAlaVal.....TyrLeuLys 154
XX
595 GCTTTCGGCATCATGAACATCGGGAAGCGTACCTACAGACATATCTTCAG 644
XX
155 LeuAspLeuLeuVal.....AsnGlyValLeuAlaLeuArgCys 167
XX
645 CTGATGGTGTATGTCTTAAACGACATCAAAATCCCAAGTTCTCATAA 694
XX
167 sLeuGluGluPheSerAlaThrAlaIleSerSerProgluProgluLeu 184
XX
695 CCTGATGAAGAAGGAGGACGCAAAACTGTCGGCAATTTCGATTCG 744
XX
184 rGlu.....CysGluValSerglyLeuLeuProleuArgProgluSer 198
XX
745 ACTTTTATTCATAAATGTTGGGGATTTTTCAGCTCCGAGCTGCTGAA 794
XX
199 SerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaIleProPh 215
XX
795 GAAATTAAGCATTCAGTGTCCAAACCTTCCTGCTGATCCGGAATCAAG 844
XX
215 eLeuThrTyrPheGlyLeuPheGluVal 224
XX
845 TGGCAGCTACTTTGGGGCTTTCAAAGTT 872
XX
seq_name: /SIDSL/gcdata/geneseq/geneseqn-embL/NA2001A.DAT:AAD15310
seq_documentation_block:
ID AAD15310 standard; cDNA; 1630 BP.
XX
AC AAD15310;
XX
DT 15-NOV-2001 (first entry)
XX
DE Murine receptor activator of NF kappaB ligand (RANKL) cDNA.
XX
KW Murine receptor activator of nuclear factor kappaB ligand; RANKL; NF;
KW tumour necrosis factor; TNF; TNF receptor associated factor; TRAF;
KW immune response; inflammatory response; graft-versus-host reaction;
KW toxic shock; sepsis; acute inflammatory reaction; bone resorption;

```

```

KW anti-apoptotic signal; therapy; immunosuppressant; anti-inflammatory; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 3..887
FT     /*tag= a
FT     /product= "Murine RANK ligand (RANKL) protein"
FT     /note= "CDS does not include start codon"
FT     /partial
XX
XX US6271349-B1.
XX
XX 07-AUG-2001.
XX
XX 17-DEC-1998; 98US-0215649.
XX
XX 23-DEC-1996; 96US-0059978.
XX
XX 07-MAR-1997; 97US-0077181.
XX
XX 14-OCT-1997; 97US-0064671.
XX
XX 23-DEC-1996; 96US-0772330.
XX
XX 07-MAR-1997; 97US-0813509.
XX
XX 22-DEC-1997; 97US-0996139.
XX
XX (IMNV ) IMMUNEX CORP.
XX
XX Dougall WC, Gallibert L;
XX
XX WPI: 2001-520313/57.
XX
XX P-PSDB: AAE08737.
XX
XX New receptor activator of NF-kappaB (RANK) polypeptides, useful for
XX regulating immune response, in screening for RANK inhibitors, or as an
XX adjunct therapy for disease characterized by neoplastic cells that
XX express RANK
XX
XX Example 7: Column 63-66; 47pp; English.
XX
XX The patent discloses novel receptor activator of nuclear factor (NF)-
XX kappaB (RANK) proteins and their corresponding DNAs. RANK is a member
XX of the tumour necrosis factor (TNF) receptor superfamily and associates
XX with TNF receptor associated factor (TRAF) 2 and 3 which are important
XX in the regulation of immune and inflammatory response. The receptors
XX are useful for regulating immune response and in screening for inhibitors
XX of these receptors. The cytoplasmic domain of RANK is used in developing
XX assays for inhibitors of signal transduction, e.g. for screening the
XX molecules that inhibit interaction of RANK with TRAF1, TRAF2, TRAF3,
XX TRAF5 and particularly TRAF6. NF-kappaB inhibition by RANK antagonists
XX are useful in ameliorating the negative effects of an inflammatory
XX response that result from triggering of RANK, e.g. in treating toxic
XX shock or sepsis, graft-versus-host reactions, acute inflammatory
XX reactions and the effects of bone resorption. RANK acts as an anti-
XX apoptotic signal and rescue the cells that express RANK from apoptosis.
XX Soluble forms of the receptor are used in vivo or in vitro based
XX screening tests for agonists or antagonists of RANK activity, as
XX antagonists of RANK-mediated NF-kappa B activation, or to inhibit
XX transduction of a signal via RANK. RANK compositions are used in the
XX development of both agonistic and antagonistic antibodies, or as an
XX adjunct therapy for disease characterised by neoplastic cells that
XX express RANK. Compounds that interfere with RANK/TRAF6 interactions
XX are useful for modulating the formation of osteoclasts from osteoclast
XX precursors and for modulating osteoclast function and activities. They
XX are used as inhibitors of diseases associated with excess bone resorption
XX and as immunosuppressants or anti-inflammatory agents. The RANK DNAs are
XX useful for the expression of recombinant proteins, as probes for analysis
XX of the presence or distribution of RANK transcripts, while the proteins
XX are useful in preparing kits for the detection of soluble RANK, or
XX monitor RANK-related activity. The present sequence is a cDNA encoding
XX murine RANK ligand (RANKL) protein.
XX
XX Sequence 1630 BP; 436 A; 355 C; 379 G; 460 T; 0 other:

```

alignment_scores:
 Quality: 106.50 Length: 276
 Ratio: 0.772 Gaps: 13
 Percent Similarity: 50.000 Percent Identity: 24.275

alignment_block:
 US-09-245-198a-2 x AAD15310 ..

Align seg 1/1 to: AAD15310 from: 1 to: 1630

```

3 SerLeuGIlyLeuAlaLeuAlaCysLeuGIlyLeuLeuValValValse 19
||||| ||||||| ||||||| |||||||
78 TCACATGTCCTGGCCCTCCCTGGGCGACCTGGCCAGCTGGCTGCAG 127
19 rLeuGIlySerTrpAlaThrLeuSerAlaGln..GluProSerGlnGlu 35
||||| ||||||| ||||||| |||||||
128 CATCCCTGTCCTGCTGACTTTCGAGCGCAGATGATCTCTACACAGA... 173
35 luleuThrAlaGluAspArg..... 41
|||||
174 ....ATATCAGAAACAGACAGACTCACTGCTTTATAGAAATCCTGAGACTC 218
42 ArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValVa 58
||||| ||||||| ||||||| |||||||
219 CATGAAACGCGATGTCGAGACTCGACTCGAGAGTGCAGACACT 268
58 lPro..... 59
|||||
269 ACCTCACTCTCGCAGAGAGATGAACACCTTTTCAGGGGGCCGTGCACA 318
60 ..PheLeuGluGlnLeuValArgProArgArg.....SerAlaProlys 73
||||| ||||||| ||||||| |||||||
319 AGGACTCTCAACACACTGTGGGGCCACAGCGCTTCTCAGGACTCTCAGCT 368
74 .....GlyArgLysAlaArgProArgAr 81
|||||
369 ATGATGGAAGGCTCATGTTGGATGTGGCCCGCAGCGCAACAGCTGAGGC 418
81 gAlaIleAlaIleHisTyrGluValHisProArgProGlyGlnAspGlyA 98
||||| ||||||| ||||||| |||||||
419 CCAGCCATTTGCACACCTCACCATCAATGCTGCCAGCATCCCATCGGGTT 468
98 lalGlnAlaGlyValAspGlyThrValSerGlyTyrGluGlu..... 111
||||| ||||||| ||||||| |||||||
469 CCCAT.....AAAGTCACTCTGCTCTCTGTTGACACAGATCGAGGC 509
112 ..ThrLysIleAsnSerSerSerProLeuArgTyrAspArgGlnIleG1 127
||||| ||||||| ||||||| |||||||
510 TGGGCCAAGATCTTAACATGACG.....TTAAGCAACGG 544
127 yGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr.....C 141
||||| ||||||| ||||||| |||||||
545 AAAAATAAGGGTTAACCAAGATGCTCTTATTACTGTACGCCCAACATT 594
141 ysgInValHisPheAspGluGlyLysAlaVal.....TyrLeuLys 154
||||| ||||||| ||||||| |||||||
595 GCTTTCGGCATCATGAACATCGGAGACGCTACACACATCATCTTCG 644
155 LeuAspLeuLeuVal.....AsnGlyValIleAlaLeuArgGly 167
||||| ||||||| ||||||| |||||||
645 CTGAGTGCTATGCTGTTAAACACAGCATCAAAATCCCAAGTTCATCA 694
167 sLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeu 184
||||| ||||||| ||||||| |||||||
695 CTTGATGAAGGAGGAGCAGAAAAAAGCTGCGGCAATTCTGAATTC 744
184 rGleu.....CysGlnValSerGlyLeuLeuProLeuArgProGlySer 198
||||| ||||||| ||||||| |||||||
745 ACTTTATTCATTAATGTTGGGGATTTTCAAGCTCGACACTGCTGTA 794
199 SerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPh 215
||||| ||||||| ||||||| |||||||
795 GAAATTAACATTCAGGTGCCAACCTTCCTCGATCCGATCCGATCAAGA 844
```

215 eleuThrTyrPheGlyLeuPheGlnVal 224
 ||||||| ||||||| ||||||| |||||||
 845 TCGCAGCTACTTGGGGCTTCAAGATT 872

seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: AAD08714

seq_documentation_block:

ID AAD08714 standard; cDNA; 1630 BP.

AC AAD08714;

DT 04-SEP-2001 (first entry)

DE Murine receptor activator of NF- χ B ligand (RANKL) cDNA.

KW Murine; receptor activator of NF- χ B; RANK; tumour necrosis factor;
 CD40; TNF receptor-associated factor; TRAF; ligand; immune response; TNF;

KW chromosome 18q22.1; RANKL; chromosome 13q14; transmembrane protein; ss.

OS Mus musculus.

FX Key Location/Qualifiers

FT CDS 3..887.

FT /*tag= a /product= "Murine RANKL protein"

FT /note= "CDS does not include start codon"

PN US6242213-B1.

PD 05-JUN-2001.

XX 22-DEC-1997; 97US-0995659.

XX 23-DEC-1996; 96US-0059978.

PR 07-MAR-1997; 97US-0077181.

PR 14-OCT-1997; 97US-0064671.

XX (TMV) IMMUNEX CORP.

XX Anderson DM;

XX WPI: 2001-407216/43.

DR P-PsDB: AAE04425.

XX New DNA molecules, useful for producing ligands (which are useful for
 regulating immune response and in screening for inhibitors of NF- χ B
 receptor activator) of the receptor activator of NF- χ B (RANK) -

XX Example 7; Column 57-60; 43pp: English.

XX The present invention relates to receptor activator of NF- χ B (RANK)

CC DNA. RANK is mapped to chromosome 18q22.1 and its ligand (RANKL) to

CC chromosome 13q14. RANK and RANKL are type 1 and type 2 transmembrane

CC proteins respectively. RANK is a member of the tumour necrosis factor

CC (TNF) superfamily and it closely resembles CD40 in the extracellular

CC region. RANK associates with TNF receptor-associated factor (TRAF) 2

CC and TRAF3. The DNA molecules are useful for producing ligands of RANK.

CC The ligands are useful for regulating immune response and in screening

CC for inhibitors of RANK. The present sequence is murine RANKL cDNA.

XX Sequence 1630 BP; 436 A; 355 C; 379 G; 460 T; 0 other;

alignment_scores:

Quality: 106.50 Length: 276
 Ratio: 0.772 Gaps: 13
 Percent Similarity: 50.000 Percent Identity: 24.275

alignment_block:

US-09-245-198a-2 x AAD08714 ..

Mon Jun 24 10:53:30 2002

Mon Jun 24 10:53:30 2002

THIS PAGE BLANK (USPTO)

OM of: US-09-245-198A-2 to: GenEmbl.* out_format : pfs
Date: Jun 22, 2002 2:25 PM
About: Results were produced by the Gencore software, version 4.5.
Copyright (c) 1993-2000 Comphen Ltd.

```

Query: US-09-245-198A-2
Query length: 225
Database: GenBank1.*
Database sequences: 1797656
Database length: 1873333701
Search time (sec): 2600.680000

```

SeqName	Strd	Len	Document
gb.ro:AF030100	+	1162.00	AF030100 Mus musculus TWAK mRNA
gb.pat:AR144047	+	1020.00	AR140407 Sequence 1 from patient
gb.pr:AF030099	+	1020.00	AF030099 Homo sapiens TWAK mRNA
gb.pat:AX201324	+	1020.00	AX201324 Sequence 3 from patient
gb.pr:AF058572	+	1020.00	AF058572 Homo sapiens Apol3/DR3
gb.pr:BC019047	+	951.00	BC019047 Sequence 1 from patient
gb.hlg:AC069459	+	602.50	AC069459 Homo sapiens, Simla
gb.ro:AL603077	-	602.50	AL603077 Mouse DNA sequence
gb.hlg:AC018676	-	545.50	AC018676 Homo sapiens DNA
gb.hlg:AC098923	-	423.50	AC098923 Rattus norvegicus
gb.ov:G62A3435	+	109.50	G62A3435 Gallus gallus mRNA
gb.pr:AL533138	+	109.50	AL533138 Human DNA sequence
gb.ba:SC12	-	108.50	SC12 Streptomyces coelicolor
gb.ba:AE004602	-	107.50	AE004602 Pseudomonas aeruginosa
gb.hlg:AC061974	+	107.00	AC061974 Homo sapiens chromo
gb.hlg:AC060789	-	107.00	AC060789 Homo sapiens chromo
gb.pat:AR156433	+	106.50	AR156433 Sequence 10 from patient
gb.pat:AR164147	+	106.50	AR164147 Sequence 10 from patient
gb.pat:AR147987	+	106.50	AR147987 Sequence 10 from patient
gb.ro:AF019048	+	106.50	AF019048 Mus musculus receptor
gb.pat:AX200995	+	106.50	AX200995 Sequence 625 from patient
gb.pat:AX267730	+	106.50	AX267730 Sequence 704 from patient
gb.pat:AX200993	+	106.50	AX200993 Sequence 623 from patient
gb.pat:AX267728	+	106.50	AX267728 Sequence 702 from patient
gb.pat:AX200996	+	106.50	AX200996 Sequence 626 from patient
gb.pat:AX267731	+	106.50	AX267731 Sequence 705 from patient
gb.hlg:AC096533	+	106.50	AC096533 Homo sapiens chromo
gb.hlg:AC093975	+	106.00	AC093975 Mus musculus chromo
gb.hlg:AC073779	+	106.00	AC073779 Mus musculus clone
gb.hlg:AC073801	+	106.00	AC073801 Mus musculus clone
gb.pr:AY070219	+	105.50	AY070219 Homo sapiens lymphoblast
gb.pr:U081YTX0XB	+	105.50	U081YTX0XB Homo sapiens lymphoblast
gb.hlg:AL451008	+	105.50	AL451008 Homo sapiens chromosome
gb.hlg:AL560009	+	105.50	AL560009 Homo sapiens chromosome
gb.pr:HSY14768	+	105.50	HSY14768 Homo sapiens DNA, cosmid
gb.pr:AP000505	+	105.50	AP000505 Homo sapiens genome
gb.hlg:AL662801	+	105.50	AL662801 Homo sapiens chromosome
gb.hlg:AL662847	+	105.50	AL662847 Homo sapiens chromosome

gb_pr:AD201G22	+	105.50	95.37	6.0e+03	184666	i	AF129756 Homo sapiens MSN
gb_pl:AP002482	+	105.50	95.23	6.1e+03	187835	i	AP002482 Oryza sativa gen
gb_da:AE005079	+	104.00	115.36	4.49.43	12797	i	AE005079 Halobacterium sp.
gb_hlg:AC061972	-	104.00	94.35	6.8e+03	164425	i	AC061972 Homo sapiens cld
gb_pr:AC079384	+	103.50	92.91	8.2e+03	180555	i	AC079384 Homo sapiens l2q
seq_name: gb_ro:AF030100							
seq_documentation_block:							
LOCUS AF030100			1168 bp	mRNA	linear		ROD 20-DEC-1997
DEFINITION Mus musculus TWEAK mRNA, partial cds.							
ACCESSION AF030100							
VERSION AF030100.1			GI:2707220				
KEYWORDS house mouse,							
SOURCE Mus musculus							
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
REFERENCE 1 (bases 1 to 1168)							
AUTHORS Chicheportliche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,							
Hession,C., Garcia,I. and Browning,J.L.							
TWEAK, a new secreted ligand in the tumor necrosis factor family							
that weakly induces apoptosis							
J Biol. Chem. 272 (51), 32401-32410 (1997)							
JOURNAL MEDLINE 96070415							
REFERENCE 2 (bases 1 to 1168)							
AUTHORS Chicheportliche,Y., Bixler,S., Tizard,R. and Browning,J.							
TITLE Direct Submission							
JOURNAL Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,							
Cambidge, MA 02142, USA							
FEATURES source Location/Qualifiers							
source 1..1168							
/organism="Mus musculus"							
/db_xref="taxon:10090"							
/cell_type="peritoneal macrophages"							
<1..679							
/note="ligand in the TNF family; secreted protein"							
/product="TWEAK"							
/codon_start=2							
/protein_id="AAC53517.1"							
/db_xref="GI:2707221"							
/translation="VLSGLAACLGLLYVSLGSWATLSAQPSOEELTAEDRRP							
PFLNPQESQDVPLEQLVRPRRSAPKRAKPAIRAIAHYEVHRRPDQAQAV							
DGVSGMEETKINSSEPLRYDRDIGETVIRAGIYYLCVHFDEGRAYVLIKDLVN							
GVALRIELEFSATASSPPQRLICVSGSLPLRPSSLRITLPMWLKAAPFLTY							
FGLFOYH"							
BASE COUNT 242 a 360 c 298 g 268 t							
ORIGIN							
alignment_scores:							
Quality: 1162.00 Length: 225							
Ratio: 5.164 Gaps: 0							
Percent Similarity: 100.000 Percent Identity: 100.000							
alignment_block:							
US-09-245-198A-2 x AF030100 ..							
Align seg 1/1 to: AF030100 from: 1 to: 1168							
1 ValLeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValVa 17							
2 GNGCGAGCGGTGGCCGTGGCCGTGCCCTTGCCCTCTCTGTGT 51							
17 lValSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnG 34							
52 GGTCAAGCTGGGGAGCTGGGCACAACGTCCTCCAGGAGCCCTTCCA 101							
34 luGluLeuThrAlaGluAspArgArgGluProProGluLeuAsnProGln 50							
102 AGGAAGCTACACAGAGGACCCTGGGAGACCCCTGAACTGAAATCCCAG 151							
51 ThrGluGluSerGlnAspValValProPheLeuGlnGluGlnLeuValAlaPr 67							

```

|||||
152 ACAGAGAAAGCCAGATGTGTACTTCTTGGAACAACACTACTCCGCC 201
67 oarqagSerAlaProLysGlyValArgLysAlaArgProArgAlaIleA 84
|||||
202 TCGAGAAAGTGTCTTAAGAGCCGGAAGCGCGCTCCGCCAGCTATTG 251
84 laalahlSTyrgLValHisProArgProGlyGlnAspGlyAlaGlnA 100
|||||
252 CAGCCATTATGAGTTTCCTCTCGGCCAGAGACGATGAGACAGCAAGCA 301
101 GlyValAspGlyThrValSerGlyTTPGluGluThrLysIleAsnSer 117
|||||
302 GGATGGATGGCAGACGTAGTGGCTGGAAAGAGACCAAAATCAACAGCT 351
117 rSerProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArg 134
|||||
352 CAGCCCTCGCGCTACGACCGCCAGATTGGGGAATTTACAGTCATCAGG 401
134 laGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlyGlySAla 150
|||||
402 CTGGGCTCTACTACTGTACTGTGACGTGACATTGATGAGGAAAGGCT 451
151 ValTyrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArgC 167
|||||
452 GTCTACTCTAAGCTGACTGTGTGAACGGTGTCTGCGCCCTGCGCTG 501
167 sleuGluGluPheSerAlaThrAlaIleSerSerProGlyProGlyLeu 184
|||||
502 CCGGAAGATTCTCAGCCACAGCAGCAAGCTCTCGGGCCCGCCAGCTCC 551
184 rGluCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeu 200
|||||
552 GTTGTGCCAGGTGTCTGGCTGTGCTGCGCCAGGAGCTTCTCCCTT 601
201 ArgIleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeu 217
|||||
602 CGGATCCGACCCCTCGGGCTCATTAAGGCTCCCGCTTCTTAAC 651
217 rTyrPheGlyLeuPheGlnValHis 225
|||||
652 CTACTTGGACTCTTTCAGACTTAC 676
seq_name: gb_pat:ARI40407
seq_documentation_block:
LOCUS ARI40407 1236 bp DNA Linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6207642.
ACCESSION ARI40407
VERSION ARI40407.1 GI:14482903
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Willey,S.R.
TITLE Member of the TNF family useful for treatment and diagnosis of
disease
JOURNAL Patent: US 6207642-A 1 27-MAR-2001;
FEATURES
SOURCE 1.1236
BASE COUNT 225 a 416 c 358 g 237 t
ORIGIN
alignment_scores:
Quality: 1020.00 Length: 224
Ratio: 4.744 Gaps: 0
Percent Similarity: 95.982 Percent Identity: 88.839
alignment_block:
US-09-245-198a-2 x ARI40407

```

```

Align seg 1/1 to: ARI40407 from: 1 to: 1236
2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValA 18
|||||
76 CTGGGGCTGGGCTGGCGCTGCGCTCTCTGCGCTCTCTGCTGCGCTGGT 125
18 lSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||
126 CAGTTTGGGAGCCGGCAGATCGCTGTCCGCCAGAGAGCTGCGCCAGAGG 175
35 lLeuThrAlaGluAspArgArgGluProProGluLeuAsnProGlnThr 51
|||||
176 AGCTGTGCAGAGGAGAGACACAGAGACCGCTCGGAATGATCCCGAGAA 225
52 GluGluSerGlnAspValAlaProPheLeuGluGluLeuValAlaProAr 68
|||||
226 GAGAAAGCCAGATCTCGCGCTTCTCTGAAACCGACTGCTGCGCTCG 275
68 garSerAlaProLysGlyValArgLysAlaArgProArgAlaIleAlaA 85
|||||
276 AAGAGTGCACCTAAGCGCGGAAAGACAGGCGGCTCGAAGAGCATCGAG 325
85 lahlSTyrgLValHisProArgProGlyGlnAspGlyAlaGlnAgl 101
|||||
326 CCATTATGAAAGTTCATCCAGCAGCTGAGACAGAGCAGCGCGCAGAT 375
102 ValAspGlyThrValSerGlyTTPGluGluThrLysIleAsnSerSer 118
|||||
376 GTGAGCGGACAGTAGTGTGCTGGAGGAAGCCAGATCAACAGCTCCAG 425
118 rProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaG 135
|||||
426 CCTGTGCGCTACACCGCCAGATCGGGAGTTATATGATCACCGCGGCTG 475
135 lLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlySAlaVal 151
|||||
476 GGCTTACTACTACCTGTACTGTGACGTGCACCTTGTATGAGGGAAGGCTG 525
152 TyrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArgCysLe 168
|||||
526 TACTGAGAGCTGACTTGTGCTGGATGTGTCTGCGCTGCGCTGCGCT 575
168 uGluGluPheSerAlaThrAlaIleSerSerProGlyProGlnLeuArg 185
|||||
576 GGAGGAATCTCAGCCACTCGCGGAGTTCCCGCGGCCCGCCAGCTCCGCC 625
185 euCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
|||||
626 TCTGCCAGGTGTGCTGCTGTGGCCCTGCGGCGGAGGCTCTCCGCGG 675
202 lLeuArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThr 218
|||||
676 ATCCGACACCTCCCTCGGGCCCATCTCAAGGCTGCGCCCTCTCTCACTA 725
218 rPheGlyLeuPheGlnValHis 225
|||||
726 CTTCGAGACTCTTCCAGGTTTAC 747
seq_name: gb_pr:AF030099
seq_documentation_block:
LOCUS AF030099 1306 bp mRNA Linear PRI 20-DEC-1997
DEFINITION Homo sapiens TWEAK mRNA, complete cds.
ACCESSION AF030099
VERSION AF030099.1 GI:2707218
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1306)
AUTHORS Chicheportiche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,

```

TITLE Hession, C., Garcia, I. and Browning, J. L.
JOURNAL that weakly induces apoptosis
MEDLINE J Biol. Chem. 272 (51), 32401-32410 (1997)
REFERENCE 2 (bases 1 to 1306)
AUTHORS Boudon, P., Hession, C., Tizard, R. and Browning, J.
TITLE Direct Submission
JOURNAL Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
Cambridge, MA 02142, USA

FEATURES
source
1. .1306
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/tissue_type="tonsil"
18. .767
/note="ligand in the TNF family; secreted protein; start
codon not verified experimentally"
/product="TNFAK"
/protein_id="AAC51923.1"
/db_xref="GI:2707219"
/translation="MAARSORRRGRGPGTALVPLALGLALACGLLAVVSL
GSRASLAQEPADPELAEEDDPSLNQTESODAPFLNLRVPRSPGRKTR
ARRATIAHYEVHPRGDDGAGVQDCTVGEARINSSPLRYNIOIEFLYTRAGL
YLYCOVHFDEGKAVYLYKLDVLYALRCLFEFSATYASSLGPOLRLOVSGLLAL
RFGSSLRITLPAHLKAAFLYFGLFQVH"

CDS
BASE COUNT 247 a 434 c 368 g 257 t
ORIGIN

alignment_scores:
Quality: 1020.00 Length: 224
Ratio: 4.744 Gaps: 0
Percent Similarity: 95.982 Percent Identity: 88.839

alignment_block:
US-09-245-198a-2 x AF030099 ..

Align seg 1/1 to: AF030099 from: 1 to: 1306

```
2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValVa 18
|||||
93 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGT 142
18 lSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||
143 CAGTTTGGGAGCGCGGCATCGTGTCCGCCAGAGCCTGCCAGAGAG 192
35 lueLeuThrAlaGluAspArgArgGluProGluLeuAsnProGlnThr 51
|||||
193 AGCTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACA 242
52 GluGluSerGlnAspValValProPheLeuGluGlnLeuValArgProAr 68
|||||
243 GAAAGAAAGCCAGATCCCTGGCCTTTCTGAAACGACTAGTTCGGCTCG 292
68 gArgSerAlaProLysGlyArgGlyAlaArgProArgAlaIleAla 85
|||||
293 CAGAGAGTCACACTAAAGCCGGAACACGGGCTCGAAGAGAGATCCGAG 342
85 lAhIstYrGluValHisProArgProGlyGlnAspGlyAlaGlnAla 101
|||||
343 CCCATTAAAGATTCAATCCAGACTGGACAGAGAGAGAGAGAGAGAGAG 392
102 ValAspGlyThrValSerGlyTrpGluGlnThrLysIleAsnSerSer 118
|||||
393 GTGAGACGGAGACAGTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 442
118 rProLeuArgTrpAspArgGlnIleGlyGluPheThrValIleArgAla 135
|||||
```

```
443 CCTCTGGCTACAAACCGCAGATCGGGAGTTATATGACACCCGGCGT 492
135 lYleuTYrTYrLeuTYrCYsgInValHisPheAspGluGlyAlaVal 151
|||||
493 GGCCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 542
152 TYrLeuTYrLeuAspLeuValAsnGlyValLeuAlaLeuArgCysLe 168
|||||
543 TACCGAAGCTGGAGCTTCTGTGTGATGTGTGTGTGTGTGTGTGTGTGT 592
168 uGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgL 185
|||||
593 GGAGGAATTCACAGCCACTGGGCGCAGATCCCTGGGCGCCAGCTCCGCC 642
185 euCYsgInValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
643 TCTGCAGAGTGTGTGGCTGTGGCCCGGCCAGGAGGCTCTCCCTCGGG 692
202 lIleArgThrLeuProTrpAlaHisIleuTYrAlaAlaProPheLeuTYr 218
|||||
693 ATCCGACACCTCCCTGGGCGCCATCTCAAGCTGCCCTTCTCACCCTA 742
218 rPheGlyLeuPheGlnValHis 225
743 CTTCGAGCTCTTCAGAGTTTCAC 764
```

seq_name: gb_pat:AX201324

seq_documentation_block:
LOCUS AX201324 1353 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 3 from Patent WO0153486.
ACCESSION AX201324
VERSION AX201324.1 GI:15391154
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 1353)
Ashkenazi, A.J., Goddard, A., Godowski, P.J., Gurney, A.L.,
Hillan, R.J., Marsters, S.A., Pan, J.J., Pittl, R.M., Roy, M.A., Smith, V.,
Stone, D.M., Watanabe, C.K. and Wood, W.I.
Compositions and methods for the treatment of tumour
Patent: WO 0153486-A 3 26-JUL-2001;
Genentech, Inc. (US)

TITLE
JOURNAL
FEATURES
source
1. .1353
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 257 a 443 c 389 g 264 t
ORIGIN

alignment_scores:
Quality: 1020.00 Length: 224
Ratio: 4.744 Gaps: 0
Percent Similarity: 95.982 Percent Identity: 88.839

alignment_block:
US-09-245-198a-2 x AX201324 ..

Align seg 1/1 to: AX201324 from: 1 to: 1353

```
2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValVa 18
|||||
133 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGT 182
18 lSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||
183 CAGTTTGGGAGCGCGGCATCGTGTCCGCCAGAGAGAGAGAGAGAGAGAG 232
35 lueLeuThrAlaGluAspArgArgGluProGluLeuAsnProGlnThr 51
|||||
```


168 uclugluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgL 185
|||||
633 GGAGGAATTTTCAGCCACTGCGGCGAGTTCCCTCGGGCCCGACGCTCCGCC 682
185 euCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
|||||
683 TCTGCAGAGTGTCTGGCTGTGGCCCTCGGCGCAGGGTCTCTCCCTCGCG 732
202 llaargThrLeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTy 218
|||||
733 ATCCGACCCCTCCCTCGGCGCATCTCAAGCGCCCTCCCTCCACCTA 782
218 rPheGlyLeuPheGlnValHis 225
|||||
783 CTTCGAGCTCTTCAGGTTCCAC 804
seq_name: gb_pat:AX180714
seq_documentation_block:
LOCUS AX180714 898 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1 from Patent WO0145730.
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 898)
AUTHORS Willey,S.R.
TITLE Tweak receptor
JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;
IMMUNEX CORPORATION (US)
FEATURES
Source location/Qualifiers
1..898
/organism="synthetic construct"
/db_xref="taxon:32630"
52..873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=11
/protein_id="CAC50004.1"
/db_xref="GI:15132571"
/translation="MATGSRSTLLAFGLCLPMLDEGATSSDRMKOIEDKIEELIS
KIYHINETARIKKIGERTSSLSASASQEPDELVAEEDDPSFINOTES
ODPAPFLNLRPRRSAPKPKRTARARALAAHEVHPKODCAQAGVDTVSAMEBA
RINSSSPLRNRQIGERTVRAGLYLCQVHDEGAAVLKLDLVDGLALRCLEE
FSATPAASSLPOLRLCOVSGLLALRPGSSLRIRTLPMALHKAAPFLTYFGLQVH"
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN
alignment_scores:
Quality: 951.00 Length: 207
Ratio: 4.803 Gaps: 0
Percent Similarity: 95.652 Percent Identity: 88.889
alignment_block:
US-09-245-198a-2 x AX180714 ..
Align seg 1/1 to: AX180714 from: 1 to: 898
19 SerLeuGlySerTrrAlaThrLeuSerAlaGlnGluProSerGlnGlu 35
|||||
250 AGTTTGGGAGCGGGGCGATCGCTGCCGCCAGGACCTCGCCAGAGAGA 299
35 uLeuThrAlaGluAspArgArgGluProProGluLeuAsnProGlnInTh 52
|||||
300 GCTGTGCGACAGAGACGACGACCGCGGACACTGATCCCGACAGACG 349
52 uGluSerGlnAspValProPheLeuGlnGluLeuValArgProArg 68
|||||
350 AAGAAAGCCAGAGATCTGCGCTTCGTAACCGACTAGTTCGGCTCGC 399

69 ArgSerAlaProLysGlyArgLysAlaArgProArgArgAlaIleAla 85
|||||
400 AGAAGTGCACCTAAAGGCCGGAACACGCGCTCGAAGACGATCGCAC 449
85 aHisTyrGlnValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 102
|||||
450 CCATTATGAAAGTTCATCCAGCAGCTGAGACAGAGCGAGGAGGAGGTG 499
102 aLaspGlyThrValSerGlyTrrGluGluThrLysIleAsnSerSer 118
|||||
500 TGGAGGAGACATGATGCTGGGAGGAGGAGGAGGAGGAGGAGGAGG 549
119 ProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAla 135
|||||
550 CTTGTGCGCTACACCGCAGATCGGGAGGATTATAGTCACCGCGGCTGG 599
135 yLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaVal 152
|||||
600 GCTTACTACCTGATGCTGACAGTGCACCTTGTATGAGGGGAGGCTGTCT 649
152 yTrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuArgCysLeu 168
|||||
650 ACCTGAAGCTGACTGCTGCTGGATGGATGGTGTCTGCGCTGCGCTG 699
169 GluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArg 185
|||||
700 GAGGAATTTTCAGCCACTGCGGCGAGTTCCCTCGGGCCCGACCTCGGCT 749
185 uCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 202
|||||
750 CTGCGCAGGCTGTGGCTGTGGCCCTCGGCGCAGGGTCTCTCCGCGGA 799
202 llaargThrLeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTy 218
|||||
800 TCCGACACCTCCCTCGGCGCATCTCAAGGCGCCCTCCCTCCACCTAC 849
219 PheGlyLeuPheGlnValHis 225
|||||
850 TTTCGAGCTCTTCAGGTTCCAC 870
seq_name: gb_pr:BC019047
seq_documentation_block:
LOCUS BC019047 1651 bp mRNA linear PRI 11-DEC-2001
DEFINITION Homo sapiens, similar to tumor necrosis factor (ligand)
superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
complete cds.
ACCESSION BC019047
VERSION BC019047.1 GI:17512138
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1651)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Louis Staudt
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgcenhgr1.nih.gov

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-422L16 is from the RPI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: PBAC3.6 This sequence is the entire insert of clone RP23-422L16.

FEATURES

Source
1. 234182
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_lib="RPI-23"
complement(84050..84131)
misc_feature
|note="Sequence from uni-directional primer reads and dgrp
big dye terminator reads only."
BASE COUNT 59310 a 56824 c 57519 g 60529 t
ORIGIN

alignment_scores:
Quality: 602.50 Length: 180
Ratio: 4.782 Gaps: 2
Percent Similarity: 70.000 Percent Identity: 68.889

alignment_block:

US-09-245-198a-2 x AL603707/rev ..

Align seg 1/1 to reverse of: AL603707 from: 1 to: 234182

```
100 ALGGLYALSPGLYTHRVASERGLYTRPGUGLUThrylasease 116
:::|||||
66609 TCAGGTGTGATGGACAGTGTGGTGGGAAGACCAAAATCAACAG 66560
116 rSerSerProLeuArGTYrAspArgInIleGlyLpHeThrValIleA 133
|||||
66559 CTCACGCCCTCTGCGCTACGACACGCCGATGGGGAATTACAGTCATCA 66510
133 rGAlaGlyLeuTYrTYrLeuTYrCys_Gln..... 142
|||||
66509 GGCGTGGGCTCTACTACTCTACTGTCAGTAAAGCCTGCGCTGCCTCAG 66460
142 ..... 142
66459 GGGCAGACGAAGGCTAAGGGAGAGGAGACTGGCAAGATGGGTGGGA 66410
142 ..... 142
66409 GAGGAGACTCTGCGACTCATGAAGAATGCCCTGCTTCATGACGG 66360
143 .....ValHispeas 146
143 AGGCAGAGCTCTGGATTGGCTCCCTCTCTCTGTCACAGGTGCACTTGA 66310
66359 pGUGlyLyAlaValTYrLeuysLleuAspLeuValasnglyValL 163
146 pGUGlyLyAlaValTYrLeuysLleuAspLeuValasnglyValL 163
66309 TGAGGGAAAGGCTGTCTACTGAAAGCTGAGCTGTGGTAAACGGTGGC 66260
163 euAlaLeuArGcysLeuGluLpHeSerAlArThraIaSerSerPro 179
|||||
66259 TGAGCCCTGGCGCTGCTGGAAGAAATTCACAGCACAGCAACACTCTCT 66210
```

```
180 GLYProGlnLeuArGcysGlnValSerGlyLeuLeuProLeuArGp 196
|||||
66209 GGGCCCCAGCTCCGTTGTGCAGGTGTGCGCTGTGGCCCTGCGCGGCC 66160
196 oglySerSerLeuArGlnIleArGhrThreuprOTPrAlaIstleuLyAlaA 213
|||||
66159 AGGGTCTCCCTTGGATCCGACCCCTCCCTGCGCTGCACTTAAAGCTG 66110
213 lArProHeuThrTYrPheGlyLeuPheGlnValHis 225
|||||
66109 CCCCCTTCACACTACTTGGACTCTTCAACTTCAC 66072
```

seq_name: gp_hcg:AC016876

seq_documentation_block:

LOCUS AC016876 190358 bp DNA linear HTG 03-SEP-2001
DEFINITION Homo sapiens clone RP11-186B7, WORKING DRAFT SEQUENCE, 16 unordered
pieces.

ACCESSION AC016876 GI:15421989

VERSION AC016876.4

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 190358)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Unpublished

2 (bases 1 to 190358)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,

Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouknighter,B.,

Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,

Cooke,P., DeRellano,K., Dewar,K., Domino,M., Donnell,L., Doyle,M.,

Ferreira,P., Fitzhugh,M., Forrest,C., Funke,R., Gage,D.,

Galagan,J., Gardy,J., Grant,G., Hayes,B., Heaford,A., Horton,L.,

Howland,J., C. Johnson,R., Jones,C., Kam,L., Karatas,A., Klein,J.,

Lehoczky,J., Lien,C., Locke,K., Macdonald,P., Marquis,N.,

McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,

Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,

Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

Testaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,

Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 3, 2001 this sequence version replaced gi:13431059.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIRB

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: 13849

Center clone name: 186.B.7

Summary Statistics

Sequencing vector: M13; W77815; 3% of reads

Sequencing vector: Plasmid; n/a; 97% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 181928 bases at least Q40

Consensus quality: 185529 bases at least Q20

Consensus quality: 187335 bases at least Q20

Insert size: 168000; agarose-fp

Insert size: 188858; sum-of-ctrls

Quality coverage: 8.1 in Q20 bases; agarose-fp

Quality coverage: 7.2 in Q20 bas

* NOTE: This is a 'working draft' sequence. It currently

* consists of 16 conligs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 20703: contig of 20703 bp in length
* 20704 20803: gap of 100 bp
* 20804 21935: contig of 1132 bp in length
* 21936 22035: gap of 100 bp
* 22036 24202: contig of 2167 bp in length
* 24203 24302: gap of 100 bp
* 24303 28199: contig of 3897 bp in length
* 28200 28299: gap of 100 bp
* 28300 35021: contig of 6722 bp in length
* 35022 35121: gap of 100 bp
* 35122 40745: contig of 5624 bp in length
* 40746 40845: gap of 100 bp
* 40846 47136: contig of 6291 bp in length
* 47137 47236: gap of 100 bp
* 47237 58980: contig of 11744 bp in length
* 58981 59080: gap of 100 bp
* 59081 68164: contig of 9084 bp in length
* 68165 68264: gap of 100 bp
* 68265 94625: contig of 26361 bp in length
* 94626 94725: gap of 100 bp
* 94726 102564: contig of 7839 bp in length
* 102565 102664: gap of 100 bp
* 102665 116764: contig of 14100 bp in length
* 116765 116864: gap of 100 bp
* 116865 136408: contig of 19544 bp in length
* 136409 136508: gap of 100 bp
* 136509 156380: contig of 19872 bp in length
* 156381 156480: gap of 100 bp
* 156481 180631: contig of 24151 bp in length
* 180632 180731: gap of 100 bp
* 180732 190358: contig of 9627 bp in length.
Location/Qualifiers
1. 190358

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RPC1-11 Human Male BAC"
1. 20703
/note="assembly_fragment
clone_end:SP6
vector_side:left"
misc_feature
20804..21935
/note="assembly_fragment"
misc_feature
22036..24202
/note="assembly_fragment"
misc_feature
24303..28199
/note="assembly_fragment"
misc_feature
28300..35021
/note="assembly_fragment"
misc_feature
35122..40745
/note="assembly_fragment"
misc_feature
40846..47136
/note="assembly_fragment"
misc_feature
47237..58980
/note="assembly_fragment"
misc_feature
59081..68164
/note="assembly_fragment"
misc_feature
68265..94625
/note="assembly_fragment"
misc_feature
94726..102564
/note="assembly_fragment"
misc_feature
102665..116764
/note="assembly_fragment"
misc_feature
116865..136408
/note="assembly_fragment"
misc_feature
136509..156380
/note="assembly_fragment"

misc_feature 156481..180631
/note="assembly_fragment"
misc_feature 180732..190358
/note="assembly_fragment
clone_end:17
vector_side:right"

BASE COUNT 45173 a 48605 c 50635 g 44441 t 1504 others
ORIGIN

alignment_scores:

Quality: 549.50 Length: 192
Ratio: 4.504 Gaps: 2
Percent Similarity: 63.542 Percent Identity: 60.417

alignment_block:

US-09-245-198a-2 x AC016876/rev ..

Align seg 1/1 to reverse of: AC016876 from: 1 to: 190358

100 AlaglyValaspGlyThrValSerGlyTrpGluGluThrIysIleAsnS 116

83375 TCAGGTGTGACGGACGACGAGTGTGCGGAGGAGCCAGAAATCAACAG 83326

116 rSerSerProLeuArGlyTrpAspArgGlnIleGlyGluPheThrValIleA 133

83325 CTCACGCCCTTGGCGCTCAACCCGACAGATCGGGAGTTTATGTCACCC 83276

133 rGAlaGlyLeuTrpTrpLeuTrpCys.Gln..... 142

83275 GGGCTGGGCTCTACTACTGCTGACTGTACGTACGTAAGCCCATCTGCTGCAT 83226

142 142

83225 GGGTAACCCAGTAAGACAGTCGCAAGGCTTTCAGAGAGTGGGGAC 83176

142 142

83175 AAGCTACAGGCGCTGGAGAGGTGAGTTGGGCTTGGGTGCGATGGATGCC 83126

142 142

83125 TGCCTGCTGAGAAATTGAAATTGAGCGAGGCGAGGCGAGGCGCTGG 83076

143ValHisPheaspGluIysAlaVal 151

83075 ACTGGGCTGTGTTCGCCACCCAGGTGCACCTTGATGATGGAGGAGGCTG 83026

151 aLTYrLeuIysLeuAspLeuValAsnGlyValIleuAlaLeuArgCys 167

83025 TCTACCTGAACCTGACCTGCTGTGGATGGTGTGCTGTGGCCCTGCGCTGC 82976

168 LeuGluGluPheSerAlaThrAlaIaSerSerProGlyProGlnLeuAr 184

82975 CTGGAGGAATTCACGACCTGCGGCGAGTTCCTCGGGGCCCGACGCTCGC 82926

184 gLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuA 201

82925 CCTCTGCCAGGTGTGTGTGTGTGGCCCTCGGGCGACGAGGTCTCTCCCTGC 82876

201 rGLeArgThrLeuProTrpAlaHisLeuIysAlaIaIaProPheLeuThr 217

82875 GGATTCGACACCTCCCTCGGGCCCATCTCAAGGCTGCCCTTCTCTCAC 82826

218 TyrPheGlyLeuPheGlnValHis 225

82825 TACTTCGACACTTTCACAGGTTAC 82802

seq_name: gb_hhg:AC098923

seq_documentation_block:

LOCUS AC098923 179030 bp DNA linear HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-154B15, *** SEQUENCING IN PROGRESS

ACCESSION

***, 68 unordered pieces.
AC098923VERSION
AC098923.4 GI:17973852KEYWORDS
HTG: HTGS PHASE1.SOURCE
Norway rat.

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.REFERENCE
1 (bases 1 to 179030)

AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbata,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,R.,
Chen,Z., Chokhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Garroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dlinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Prantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Homi,F., Howard,S., Huber,J., Huiyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Jondah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korrah,J.,
Kovar,C., Kratovic,J., Kuresh,A., Landy,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Lousaged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheswari,M., Mapa,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabhat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,M., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkenko,S.,
Ogulu,M., Okunolu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojokokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtai,N.,
Slason,I., Sodergren,E., Sonake,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Uemari,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Welnsstock,G. and Gldbs,R.

Unpublished

2 (bases 1 to 179030)

Direct Submission

Submitted

Submitted (06-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17064677.

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: G10K

Center clone name: CH230-154B15

----- Summary Statistics

Assembly Program: Phrap; version 0.990329first call to

flandphraplist

Consensus quality: 137096 bases at least Q40

Consensus quality: 145080 bases at least Q30

Consensus quality: 152950 bases at least Q20

Estimated insert size: 138980; sum-of-contrigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.2x in Q20 bases; sum-of-contrigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a "working draft" sequence. It currently
consists of 68 contrigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contrigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1	7467:	contig of 7467 bp in length
7468	7567:	gap of unknown length
7568	12115:	contig of 4548 bp in length
12116	12215:	gap of unknown length
12216	20313:	contig of 8098 bp in length
20314	20413:	gap of unknown length
20414	25589:	contig of 5176 bp in length
25590	25689:	gap of unknown length
25690	31434:	contig of 5745 bp in length
31435	31534:	gap of unknown length
31535	37595:	contig of 6061 bp in length
37596	37695:	gap of unknown length
37696	42172:	contig of 4477 bp in length
42173	42272:	gap of unknown length
42273	46234:	contig of 3962 bp in length
46235	46334:	gap of unknown length
46335	50605:	contig of 4271 bp in length
50606	50705:	gap of unknown length
50706	55202:	contig of 4497 bp in length
55203	55302:	gap of unknown length
55303	58407:	contig of 3105 bp in length
58408	58507:	gap of unknown length
58508	61286:	contig of 2779 bp in length
61287	61386:	gap of unknown length
61387	64467:	contig of 3081 bp in length
64468	64567:	gap of unknown length
64568	68498:	contig of 3931 bp in length
68499	68598:	gap of unknown length
68599	73449:	contig of 4851 bp in length
73450	73549:	gap of unknown length
73550	77402:	contig of 3853 bp in length
77403	77502:	gap of unknown length
77503	81027:	contig of 3525 bp in length
81028	81127:	gap of unknown length
81128	84144:	contig of 3017 bp in length
84145	84244:	gap of unknown length
84245	86780:	contig of 2436 bp in length
86781	86780:	gap of unknown length
86781	89210:	contig of 2430 bp in length
89211	89310:	gap of unknown length
89311	92339:	contig of 3029 bp in length
92340	92439:	gap of unknown length
92440	94951:	contig of 2512 bp in length
94952	95051:	gap of unknown length
95052	98586:	contig of 3335 bp in length
98587	98686:	gap of unknown length
98687	100597:	contig of 1911 bp in length
100598	100697:	gap of unknown length
100698	102766:	contig of 2069 bp in length
102767	102866:	gap of unknown length
102867	105237:	contig of 2371 bp in length
105238	105337:	gap of unknown length
105338	107305:	contig of 1968 bp in length
107306	107405:	gap of unknown length
107406	109057:	contig of 1652 bp in length
109058	109157:	gap of unknown length
109158	111454:	contig of 2297 bp in length
111455	111554:	gap of unknown length
111555	113704:	contig of 2150 bp in length
113705	113804:	gap of unknown length

```

* 11805 115933: contig of 2129 bp in length
* 115934 116033: gap of unknown length
* 116034 118193: contig of 2160 bp in length
* 118194 118293: gap of unknown length
* 118294 120414: contig of 2121 bp in length
* 120415 120514: gap of unknown length
* 120515 122311: contig of 1797 bp in length
* 122312 122411: gap of unknown length
* 122412 125628: contig of 3217 bp in length
* 125629 125728: gap of unknown length
* 125729 128345: contig of 2817 bp in length
* 128346 128645: gap of unknown length
* 128646 130563: contig of 1918 bp in length
* 130564 130663: gap of unknown length
* 130664 132945: contig of 2282 bp in length
* 132946 133045: gap of unknown length
* 133046 134356: contig of 1511 bp in length
* 134357 134656: gap of unknown length
* 134657 136478: contig of 1822 bp in length
* 136479 136578: gap of unknown length
* 136579 138159: contig of 1581 bp in length
* 138160 138259: gap of unknown length
* 138260 140456: contig of 2197 bp in length
* 140457 140556: gap of unknown length
* 140557 142544: contig of 1988 bp in length
* 142545 142644: gap of unknown length
* 142645 143784: contig of 1140 bp in length
* 143785 143884: gap of unknown length
* 143885 145304: contig of 1420 bp in length
* 145305 145404: gap of unknown length
* 145405 147391: contig of 1987 bp in length
* 147392 147491: gap of unknown length
* 147492 148908: contig of 1417 bp in length
* 148909 149008: gap of unknown length
* 149009 150147: contig of 1139 bp in length
* 150148 150247: gap of unknown length
* 150248 151752: contig of 1505 bp in length
* 151753 151852: gap of unknown length
* 151853 153441: contig of 1589 bp in length
* 153442 153541: gap of unknown length
* 153542 154953: contig of 1412 bp in length
* 154954 155053: gap of unknown length
* 155054 156244: contig of 1191 bp in length
* 156245 156344: gap of unknown length
* 156345 158095: contig of 1751 bp in length
* 158096 158195: gap of unknown length
* 158196 159309: contig of 1114 bp in length
* 159310 159409: gap of unknown length
* 159410 160682: contig of 1273 bp in length
* 160683 160782: gap of unknown length
* 160783 162602: contig of 1820 bp in length
* 162603 162702: gap of unknown length

```

alignment_scores:

Quality: 425.50 Length: 176
 Ratio: 4.052 Gaps: 2
 Percent Similarity: 59.659 Percent Identity: 55.114

alignment_block:

US-09-245-198a-2 x AC098923/rev ..

Align seg 1/1 to reverse of: AC098923 from: 1 to: 179030

```

101 GtYAlaspGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSe 117
|||||
26419 GGTGTGATGGACAGTGTAGTGGTGGAGAGACCAAAATCAACAGCTC 26370
|||||
117 rseProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArg 134
|||||
26369 CAGCCCTCTGCGCTATGACCCCAATTGGGGAAATTTCGCTCATCAGG 26320
|||||
134 lagIlyLeuTyrTyrLeuTyrCys..... 141
|||||

```

```

26319 CTGGGCTCTACTACCTGACTGTCAGTAAGCCCTGGGCTCCATGGGTAG 26270
141 ..... 141
26269 AGCCATGGCTAAGGAGAGAGAGCTTGGCAGGAAGTGGTGGAGTGGA 26220
141 ..... 141
26219 GAACCTGGGTTTCATGAAGAAGATGCTGATTTCCATGAGGCGACGA 26170
142 .....GlnValHisPheAspGlu 147
26169 GAGCTGTGATTTTGTCTGTCTCTCTGCTCCAGGTCACCTTGATGAG 26120
148 GlyLysAlaValIlyrLeuLysLeuAspLeuValAsnGlyValIleAla 164
26119 GGGAAAGGACACTGTACTGAAGCTGACTGTGTGTAAGTGTGCTGC 26070
164 aLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerProGlyP 181
26069 CCTGCACGCTCGTGGAAAAATTATCAGCCACAGTACCGAGATCTCTGGGC 26020
181 roGlnLeuArgLeuCysGlnVal_SerGlyLeuLeuProLeuArgProG 197
26019 CCCAGATCCGTTTGGCCAGCCGCTGTGCTGTGGATCCGCGCCGCG 25970
197 yseSerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaIaP 214
25969 GTCCTCACTTTCAGACCCGCCACCCCTTAACCTCCCTTAAGCCTCCGC 25920
214 ropheLeuThrTyrPheGlyLeuPhe 222
25919 CCGTCCATTCCTCCCGCGGACTCTTT 25894

```

seq_name: gb_ov:GGA243435

seq_documentation_block:

LOCUS GGA243435 892 bp mRNA linear VRT 28-JUN-2000

DEFINITION Gallus gallus mRNA for putative CD154 (CD40 ligand).

ACCESSION AJ243435

VERSION AJ243435.1 GI:8977828

KEYWORDS CD154 gene.

SOURCE chicken.

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 892)

AUTHORS Tregaskes,C.A., Young,J.R. and Burnside,J.

TITLE cloning of a putative chicken CD40 ligand

JOURNAL unpublished

REFERENCE 2 (bases 1 to 892)

AUTHORS Tregaskes,C.A.

TITLE Direct Submision

JOURNAL Submitted (24-JUN-1999) Tregaskes C.A., Immunopathology, Institute for Animal Health, High St, Compton, Nr Newbury, Berkshire, UNITED KINGDOM

FEATURES

source

```

1..892 Location/Qualifiers
  /organism="Gallus gallus"
  /strain="white leghorn Line 0"
  /db_xref="taxon:9031"
  /clone="chcd4012.2"
  /tissue_type="spleen"
  /gene="CD154"
  /gene="CD154"
  /codon_start=1
  /product="putative CD154 (CD40 ligand)"
  /protein_id="CAB95748.1"
  /db_xref="GI:8977829"
  /translation="MNEAYSPAAPRRPMGSTSPSTMKMFQFLSVFNVQITGVLFCL

```


/note="MER53 repeat: matches 1. .186 of consensus"
3255. .3292
repeat_region
/note="19 copies 2 mer tt 92% conserved"
3878. .3954
repeat_region
/note="L2 repeat: matches 2164. .2237 of consensus"
4113. .4282
repeat_region
/note="MER5B repeat: matches 4. .174 of consensus"
4285. .4436
repeat_region
/note="MER5A repeat: matches 1. .174 of consensus"
5142. .5196
repeat_region
/note="MIR repeat: matches 201. .255 of consensus"
5206. .5305
repeat_region
/note="MER5A repeat: matches 9. .112 of consensus"
5298. .5331
repeat_region
/note="MER5A repeat: matches 43. .75 of consensus"
5442. .5744
repeat_region
/note="AluSc repeat: matches 1. .305 of consensus"
5965. .6254
repeat_region
/note="L2 repeat: matches 2423. .2707 of consensus"
6560. .6856
repeat_region
/note="AluJb repeat: matches 2. .300 of consensus"
complement(6576. .7026)
misc_feature
/note="match: GSS: Em:B63556"
7367. .7776
misc_feature
/note="match: GSS: Em:B32322"
7640. .7672
repeat_region
/note="MER5A repeat: matches 43. .75 of consensus"
7665. .7771
repeat_region
/note="MER5A repeat: matches 4. .112 of consensus"
8989. .9080
repeat_region
/note="MIR repeat: matches 51. .142 of consensus"
9189. .9701
misc_feature
/note="match: GSS: Em:AQ760102"
9190. .9605
misc_feature
/note="match: GSS: Em:B90415"
9572. .9811
repeat_region
/note="MIR repeat: matches 3. .247 of consensus"
complement(10614. .11080)
misc_feature
/note="match: GSS: Em:AQ571033"
12183. .12226
repeat_region
/note="Alu repeat: matches 253. .298 of consensus"
12235. .12348
repeat_region
/note="L2 repeat: matches 2636. .2749 of consensus"
13889. .13938
repeat_region
/note="25 copies 2 mer tt 80% conserved"
14485. .15124
misc_feature
/note="CpG Island"
/evidence="not_experimental"
14763. .14880
repeat_region
/note="59 copies 2 mer cc 62% conserved"
15224. .15285
repeat_region
/note="31 copies 2 mer gg 72% conserved"
complement(15328. .15962)
misc_feature
/note="match: GSS: Em:AQ631047"
15510. .15943
misc_feature
/note="match: GSS: Em:AQ126356"
15744. .15771
repeat_region
/note="14 copies 2 mer tg 89% conserved"
15972. .16456
misc_feature
/note="match: GSS: Em:AQ418183"
15991. .16438
misc_feature
/note="match: GSS: Em:AQ815493"
complement(16128. .16449)
misc_feature
/note="match: STS: Em:HS015X65"
16198. .16239
repeat_region
/note="21 copies 2 mer gt 97% conserved"
17613. .18726
repeat_region
/note="LIMC3 repeat: matches 6631. .7735 of consensus"
18812. .18936
repeat_region
/note="LIMC4 repeat: matches 6487. .6609 of consensus"
21659. .22737
repeat_region
/note="TIGER1 repeat: matches 1324. .2412 of consensus"
22738. .23149

/note="MER57A repeat: matches 1. .434 of consensus"
23150. .23658
repeat_region
/note="TIGER1 repeat: matches 774. .1326 of consensus"
23655. .24186
repeat_region
/note="TIGER1 repeat: matches 1. .547 of consensus"
24920. .25222
repeat_region
/note="AluY repeat: matches 1. .303 of consensus"
25712. .25869
repeat_region
/note="MER97c repeat: matches 16. .179 of consensus"
26054. .26081
repeat_region
/note="14 copies 2 mer tt 100% conserved"
27311. .27518
misc_feature
/note="match: STS: Em:G04517"
27350. .27802
misc_feature
/note="match: GSS: Em:AQ614744"
27427. .27633
repeat_region
/note="MIR repeat: matches 2. .204 of consensus"
27895. .28024
repeat_region
/note="MIR repeat: matches 111. .246 of consensus"
29211. .29252
repeat_region
/note="21 copies 2 mer at 83% conserved"
29268. .29531
repeat_region
/note="Cheshire repeat: matches 2216. .2464 of consensus"
29499. .29926
misc_feature
/note="match: GSS: Em:AQ617879"
29533. .29681
repeat_region
/note="MUTID repeat: matches 1. .154 of consensus"
29820. .30127
repeat_region
/note="MLTIE repeat: matches 255. .560 of consensus"
30142. .30804
repeat_region
/note="Cheshire repeat: matches 1528. .2173 of consensus"
30805. .31116
repeat_region
/note="AluX repeat: matches 11. .312 of consensus"
31117. .31175
repeat_region
/note="Cheshire repeat: matches 1469. .1528 of consensus"
31187. .32501
repeat_region
/note="Cheshire repeat: matches 1. .1360 of consensus"
33567. .33766
repeat_region
/note="MIR repeat: matches 48. .261 of consensus"
34845. .35149
repeat_region
/note="AluSq repeat: matches 1. .303 of consensus"
37351. .37408
repeat_region
/note="29 copies 2 mer aa 72% conserved"
37423. .37571
repeat_region
/note="MER5A repeat: matches 2. .188 of consensus"
37565. .37656
repeat_region
/note="MER5A repeat: matches 1. .105 of consensus"
37723. .38025
repeat_region
/note="AluX repeat: matches 8. .308 of consensus"
complement(37990. .38488)
misc_feature
/note="match: STS: Em:LI18341"
38062. .38113
repeat_region
/note="26 copies 2 mer ac 92% conserved"
complement(38539. .39048)
misc_feature
/note="match: GSS: Em:AQ570301"
39057. .39524
misc_feature
/note="match: GSS: Em:AQ778077"
39069. .39496
misc_feature
/note="match: GSS: Em:AQ418987"
39334. .39389
repeat_region
/note="MIR repeat: matches 90. .145 of consensus"
40906. .41237
repeat_region
/note="LIME1 repeat: matches 5817. .6122 of consensus"
41485. .41800
repeat_region
/note="AluY repeat: matches 1. .310 of consensus"
41802. .41839

alignment_scores:
Quality: 109.50 Length: 221
Ratio: 1.117 Gaps: 12
Percent Similarity: 44.344 Percent Identity: 24.887
alignment_block:

US-09-245-198a-2 x AL353138/rev ..

Align seg 1/1 to reverse of: AL353138 from: 1 to: 108967

```

42 ArgGluPro...ProGluLeuAsnProGlnThrGluGluSerGlnAsp 57
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14628 CGGAGCCGGACACCGCTCTTTCTTCACCCGCGCCGACGCTGAC... 14581
57 lValProPheLeuGluGlnLeuValArgProArgSerAlaPro... 72
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14580 .....ATCACCTTCGGCTCAGCGCCGCTCAACCGCTTC 14547
73 ....LysGlyArgGlyAlaArgProArgAlaIleAlaIleHisTyr 87
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14546 CAGCCCGAGCGGCGCTGAGCGCTGCTGCTCC..... 14512
88 GluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly..... 101
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14511 .....CAGCCACAGCTGCGCAAGATCGAGCGCGCTGCGACGACGCC 14468
102 .....ValAspGlyThrValSerG 108
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14467 TTTTCCCGCTGGACGACGTCACAAAGCCAGCTGTATTCGACTTACACACC 14418
108 lYrTrpGlu.GluThrLysIleAsnSerSerSerPro..... 119
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14417 CCTGGAGGTATGCGAGGCTGATCTCTCTCCCAACCGGCGACAATAT 14368
120 .....LeuArgTyrAspArgGlnIleGlyGluP 129
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14367 TTGTCTCCAGGCTGAGCGGATCTCTGCAAGCCAGCGCGTCTGGANA 14318
129 heThrValIleArgAlaGlyLeuTyrTyrLeuTyr.....CysGln 142
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14317 TGAAGGTGAACCAAGGCTATATTAATTAATTTCCATCTGCTGTG 14268
143 ValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuVal 159
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14267 TTGGAGAAATCAGAGTAAAGAAATGCTG.....CTGATGATATTAAT 14227
159 lAsnGlyValIleuAlaLeuValGlyLeuGluPheSerAlaThrAla 176
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14226 TTCA.....GTGCTAAGAGATCTCAGCGAGGCTGCCG 14195
176 lAsnSer.....ProGlyProGln 182
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14194 TTCCAATTAAGACAACAGCAGGCTGCAAGTGAATGCCAACAATAA 14145
183 LeuArgLeuGlyGlnValSerGlyLeuProLeuArgProGly..... 197
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14144 TTAAAGTATGCTCTTCTGATGCTTTCGCCAATTAAGAGCGACGACT 14095
198 .....SerSerLeu 201
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14094 TCATTTCTTGCCCTAAACAAGCAAGAAATAATGCAAGAGTCTCATCTTA 14045
201 rglLeuArgThr 204
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
14044 GACTCAGAGAC 14034

```

seq_name: gb_ba:SC1C2

seq_documentation_block:

LOCUS SC1C2 42210 bp DNA linear BCT 15-JAN-1999
 DEFINITION Streptomyces coelicolor cosmid 1C2.
 ACCESSION AL031124
 VERSION AL031124.1 GI:3355667
 KEYWORDS 3-isopropylmalate dehydratase large subunit; 3-isopropylmalate
 dehydratase small subunit; 3-isopropylmalate dehydrogenase;
 branched-chain amino acid aminotransferase; carboxyl transferase;
 delta-1-pyrroline-5-carboxylate dehydrogenase; glx; glutamyl-tRNA
 synthetase; histone-like DNA binding protein; hydrolase; lve;
 leuB; leuc; leuP; lyase; secreted lyase; transfer-RNA-Gln;

SOURCE

Transfer-RNA-Glu: ureaB; urease alpha subunit; urease beta and gamma subunits; ureC.
 Streptomyces coelicolor A3(2).
 Streptomyces coelicolor A3(2)

ORGANISM

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 1 (bases 1 to 42210)
 Murphy, L. and Harris, D.
 Unpublished
 2 (bases 1 to 42210)
 Parkhill, J., Barrell, B. G. and Rajandream, M. A.

REFERENCE

Submitted (23-JUL-1998) Streptomyces coelicolor sequencing project,
 Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire
 CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
 David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
 Colney, Norwich, Norfolk NR4 7UH, UK
 3 (bases 1 to 42210)
 Redenbach, M., Kleiser, H. M., Denapate, D., Eichner, A., Cullum, J.,
 Kinash, H. and Hopwood, D. A.
 A set of ordered cosmids and a detailed genetic and physical map
 for the 8 Mb Streptomyces coelicolor A3(2) chromosome
 Mol. Microbiol. 21 (1), 77-96 (1996)
 97000351

TITLE

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

JOURNAL

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

MEDLINE

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

COMMENT

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

NOTES

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

FEATURES

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

source

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

gene

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

CDS

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

FEATURES

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

source

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

gene

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

CDS

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

FEATURES

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

source

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

gene

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

CDS

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

FEATURES

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

source

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

gene

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).

CDS

Streptomyces coelicolor sequencing at The Sanger Centre is funded
 by the BBSRC.
 Details of S. coelicolor sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are
 numbered using the following system eg SC7B7.01c. SC (S.
 coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
 strand).


```
(563 aa), fasta scores; opt: 1273 z-score: 1730.4 E(): 0,
48.5% identity in 408 aa overlap. Contains PS00687 and
PS00070 Aldehyde dehydrogenases glutamic acid active site
and cysteine active site and Pfam match to entry PF00171
aldehyd, Aldehyde dehydrogenases, score 114.70, E-value
1.8e-30"
/codon_start=2
/transl_table=11
/product="delta-1-pyrroline-5-carboxylate dehydrogenase"
/protein_id="CA19968.1"
/db_xref="GI:3355668"
/db_xref="SPTREMBL:O86502"
/translation="IDSPCLIDFMFNVHYARNILAEQPPANSGVNNRMDHPLEG
FYATPPFNESAIAANLPTAPALMGVNVWKPSPOTHAAILMOLLEPAGLPKVIN
LYTGGIAVSEVALHEHRLAGIHFGSTKTPHMKTNNTKERTYRIVAGETGK
DLVNAHPSADRAVLKATLRGAEEYOGKCSRTSKRYTPASITNDSFKEEFAEYVL
TMGDVDTLSNFIAVVIDERSFAKNKAIDRAKEDERTIVAGSYDSDVAGYFRTVAV
ECTDPENEVEFTEYEGPLAVHYDSDADAYDAMLQHEVSVDYALTSVISNDRAA
AAVTMEKRLTYAGNFEYINDKSTGAVVGQOPFGGGRASGTNDKAGADPNLMRTLTIRAI
KETLVAPDTYTPHMG"
1..3807
/feature="true overlap with cosmid 8D9"
2..1189
/feature="SC1C2.01"
/feature="Pfam match to entry PF00171 aldehyd, Aldehyde
dehydrogenases, score 114.70, E-value 1.8e-30"
464..487
/feature="SC1C2.01"
/feature="PS00687 Aldehyde dehydrogenases glutamic acid
active site"
548..583
/feature="SC1C2.01"
/feature="PS00070 Aldehyde dehydrogenases cysteine active
site"
1246..1283
/feature="hairpin_loop with 18bp stem"
1522..1527
/feature="possible RBS upstream of SC1C2.02"
1537..1668
/feature="SC1C2.02"
1537..1668
/feature="SC1C2.02"
1537..1668
/feature="SC1C2.02"
/feature="SC1C2.02, questionable ORF, len: 43 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein SC1C2.02"
/protein_id="CA19969.1"
/db_xref="GI:3355669"
/db_xref="SPTREMBL:O86503"
/translation="MAETTVRRVRHLRSTSDSRKNAALQRLDRDNGATGH"
1744..2787
/feature="leuB"
1744..2787
/feature="leuB"
/feature="leuB"
/feature="SC1C2.03, leuB, probable 3-isopropylmalate
dehydrogenase, len: 347 aa; similar to e.g. LEU3 CORGL
3-isopropylmalate dehydrogenase (EC 1.1.1.85) (340 aa),
fasta scores; opt: 1450 z-score: 1318.1 E(): 0, 68.08
identity in 338 aa overlap. Contains PS00470 Isocitrate
and isopropylmalate dehydrogenases signature and Pfam
match to entry PF00180 isodh, Isocitrate and
isopropylmalate dehydrogenases, score 356.60, E-value
1.3e-105"
/codon_start=1
/transl_table=11
/product="3-isopropylmalate dehydrogenase"
/protein_id="CA19970.1"
/db_xref="GI:3355670"
/db_xref="SPTREMBL:O86504"
/translation="MSRSLSLIAVTPGDGIGOEYVAGRLKVLISAVLPDVKLTKEKDF
GARRVHATSETLTADUDALKAHDAILLGAIGSPSPSVLENGFLTKLRPAFDHVN
LRPSKLPGVATPLAGQPEIDVYVREGTEGPTGCGTIRKTEHEVATVESVNTAY
GVERVVRDAFARAQARPRKLTLLVHKNNVLTFAGHILMTIFNKVAAEYDEVTYDLHV
```

```
DAATIFLVTPARFDVITVDNLGDIITDILAAVSGCIGVAASGNINPSGDEPSPNEP
VHGSADIDGCKADPTATVLSALLRLHLYGEDEARAIEDAVSADLCGRGLPARST
SEIGDRLARVAG"
1759..2766
/feature="leuB"
/feature="Pfam match to entry PF00180 isodh, Isocitrate and
isopropylmalate dehydrogenases, score 356.60, E-value
1.3e-105"
2458..2517
/feature="leuB"
/feature="PS00470 Isocitrate and isopropylmalate
dehydrogenases signature"
3038..3042
/feature="possible RBS upstream of lIve"
3052..4140
/feature="lIve"
/feature="lIve"
/feature="lIve"
/feature="lIve"
/feature="SC1C2.04, lIve, probable branched-chain amino acid
aminotransferase, len: 362 aa; similar to many, e.g.
lIve_BACSU putative branched-chain amino acid
aminotransferase (362 aa), fasta scores; opt: 1047
z-score: 1401.5 E(): 0, 44.7% identity in 360 aa overlap.
Contains PS00770 Aminotransferases Class-IV signature"
/codon_start=1
/transl_table=11
/product="branched-chain amino acid aminotransferase"
/protein_id="CA19971.1"
/db_xref="GI:3355671"
/db_xref="SPTREMBL:O86505"
/translation="MTTPRIETLKPSPHPLSDSERAILANPFGRHFTDHMTYIKTE
GRWHDGOLVPAAPLPLSDPATVILHYAOEIFPGLKRYRRDPSVATFREKNGARPOA
SSRRIGMPELPVDPTFEACDIALVAODEKVPARHGESELYLRPMATVGVAPRAN
EYLFTYASPAAGATPPGKVPYSIYSESRVRAVPGMGKDATGCTATASLLAQRAA
AKGCDVCTLDAIERKAVEELGGMNLYFYGNKIVTPSLTGSILEVTRDSLLTVARD
LGYEADEGRVSVDQWQDSENGTLTEVFACGKAAYITPVGTWRAGAWQDSGETGE
VTQRLDALDLDIQRTVADPHGMHTLA"
3763..3852
/feature="lIve"
/feature="PS00770 Aminotransferases class-IV signature"
complement(4214..5644)
/feature="SC1C2.05c"
complement(4214..5644)
/feature="SC1C2.05c"
alignment_scores:
Quality: 108.50 Length: 225
Ratio: 0.882 Gaps: 13
Percent Similarity: 54.667 Percent Identity: 29.333
alignment_block:
US-09-245-198a-2 x SC1C2/rev ..
Align seg 1/1 to reverse of: SC1C2 from: 1 to: 42210
19 SerleuglyserTPAlathrleuSerAlaIngluProserGingluG1 35
:::|||||::: |||:::|||||::: :::::|||||:::
9451 GCGCTGGCGCGGAGCGCGCCCTGCTCACTCCAGGCTGGCGACACCA 9402
GCGCTGGCGCGGAGCGCGCCCTGCTCACTCCAGGCTGGCGACACCA 9402
35 u.....leuThrAlaGluAspArgGluPro..ProGluLeuAsnPro 49
: ||| ||||| ::||| ||| ||| |||
9401 GTTCGCTCTTGAGCGCAAGTAACAGAGGTGGCGCTGCTCATGGCG 9352
GTCGCTCTTGAGCGCGCAAGTAACAGAGGTGGCGCTGCTCATGGCG 9352
50 GlnThrGluSerGlnAspValValProhelaengluGlnLeu.ValA 66
:::|::| ||| ::||| |||
9351 ACCTCGCGCGCGCGCGCCGCGCATGCTGAGCTTC.....TCAGGCC 9311
ACCTCGCGCGCGCGCGCCGCGCATGCTGAGCTTC.....TCAGGCC 9311
66 rGPProArgSerAlaProLysGlyArgLysAlaArgProArgArgAla 82
|| |||||:::|||| |:::||||| ||| |||
9310 GCGCTGGCGCATCTCCATGCGCGCGCGACAGACGCTCGCGCGCGCG 9261
GCGCTGGCGCATCTCCATGCGCGCGCGACAGACGCTCGCGCGCGCGCG 9261
83 .....IleAlaAlaHisTyrGluValHisProArgProGlyGlnAspG1 97
:::|||||::: ||| ||||| |||
```


THIS PAGE BLANK (USPTO)

FEATURES	source	Location/Qualifiers
CDS		1..1168 /organism="Mus musculus" /db_xref="taxon:10090" /cell_type="peritoneal macrophages" <1..679 /note="ligand in the TNF family; secreted protein" /codon_start=2 /product="TWEAK" /protein_id="AAC5517.1" /db_xref="gi:2707221" /translation="VLSGLALACGLLVVSLGSMTLSAOEPSOELTAEDRREF PELNPQTESQDVVPEFLQYLPRPSRAKGRAPRRRIIAHYEVHPGDDGAQGV DGTSGMGEVETINSSPLRVDROJGEFVITAGIYYLCOVHPDEGRKYVTKLLIV GVALARCLEEFSATRAASSPQQLRICOVSGLLPLRPGSSLRILPLMAHLAAPELTY FGLRVNH"
BASE COUNT	242 a 360 c 298 g 268 t	
ORIGIN		
Query Match	100.0%;	Score 1166; DB 10; Length 1168;
Best Local Similarity	100.0%;	Pred. No. 1.1e-295;
Matches 1168; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
OY	1	ggtgtgaagccctgggagccctggcgcttggcccttgccctgcctgcctgctgtgtgtgaagcct 60
DB	1	ggtgctgaagccctgggagccctggcgcttggccctgcctgctgcctgcctgtgtgtgtgaagcct 60
OY	61	gggagagctgggcaagctgtctgcgccaagagcctctcagaagagctgacagcaagga 120
DB	61	ggggagagctgggcaagcgtgtctgtcccgagagccttctcagagagagctgacagcagaga 120
OY	121	ccgagcggaagccccctgaactgaatccccaagagaagaagccagatgtgtactctt 180
DB	121	ccgagcgagagccccctgaactgaatccccaagagaagaagccagatgtgtactctt 180
OY	181	cttggacaacactagtcgagcctcgagaaagtctcctaagggcggaagcgcgagctcg 240
DB	181	cttggacaacactagtcgagcctcgagaaagtctcctaagggcggaagcgcgagctcg 240
OY	241	ccgagcctatcgagccactatgaagttcaatccctcgccagagaacagatgtgagcaaac 300
DB	241	ccgagcctatcgagccactatgaagttcaatccctcgccagagaacagatgtgagcaaac 300
OY	301	agggtgtgagatgggagcagtgagtgctgtgggaagagacccaataatcaacagctccagccctc 360
DB	301	agcgtgtgagatgggagcagtgagtgctgtgggaagagacccaataatcaacagctccagccctc 360
OY	361	ggcgctgagacccgccaagatgtgggaattacaagttcaagggcttggtctactactctga 420
DB	361	ggcgctgagacccgccaagatgtgggaattacaagttcaagggcttggtctactactctga 420
OY	421	ctgtcaagtgtgaccttgatgaaggaagaagcgtgtcacaacggaagctggaactgtgtgta 480
DB	421	ctgtcaagtgtgaccttgatgaaggaagaagcgtgtcacaacggaagctggaactgtgtgta 480
OY	481	cgggtgtgtgagccctgtgcctctcggaagaattctcaagccacagcagaagaagctctctcgg 540
DB	481	cgggtgtgtgagccctgtgcctctcggaagaattctcaagccacagcagaagaagctctctcgg 540
OY	541	gcccagagctccgcttgtgtgcgaagtgtctgtggctgttcgcgctcgaggaagctctcct 600
DB	541	gcccagagctccgcttgtgtgcgaagtgtctgtggctgttcgcgctcgaggaagctctcct 600
OY	601	tggagatcgagaccctcccccctgggctcaatcttaaggttgccccctctctaactactttg 660
DB	601	tggagatcgagaccctcccccctgggctcaatcttaaggttgccccctctctaactactttg 660
OY	661	actcttcaagttaacttgggggagcctgtctccccaagatctcttaacttcccttgagtc 720
DB	661	actcttcaagttaacttgggggagcctgtctccccaagatctcttaacttcccttgagtc 720
OY	721	caggagcatcaccaaacctccctaccccaccccacatctctcaacccccctgcgtactctt 780

DB	721	CAGAGCATCAACACAGCTCCCTACCCACCCCACTCTCCACCCCTGCTGCTCTT	780
OY	781	ggtccagtcctgctctccccaagaagcagagctgttcaaatgttccatcca	840
DB	781	GGTCCAGTCTGCTCTCTCTCAAAAGGACGACGAGCTGTTCACATGTTCCATTCCACA	840
OY	841	gaagatctctgctctcttcaactccatcccaacaaataatccactactagctcc	900
DB	841	GACGATCTCTGCTCTTCTTAACATCCATCCATCCACCAACATATCCACTACCTAGCTCC	900
OY	901	caaaagccctactatccctgactcccccacccactcaaccgagcaagtgctttagct	960
DB	901	CAAAAGCCCTACTATTATCCCTGACTCCCCACCACTACCGACACAGCTGTTATTTAGCT	960
OY	961	ttgtgaccagcagctgagatgagctgagctggtgagcaagaagccagaagacctggagc	1020
DB	961	TTGTGCACACAGCAGCTAGATGAGCTGAGCTGAGCTGTGGTGACAGCAAGCAGAACTGGGAC	1020
OY	1021	taggccaagaagttcccaactgtaggggggaagagctgtagggcaagctctccctgagacc	1080
DB	1021	TAGGCCAAGAAGTCCCAACTGTGAGGGGGAAGAGCTGGGGCAACACTCTCTCTGTGATCC	1080
OY	1081	ctgtgagctttagaagaagatactatttattatattgtgtgcaaaatgttaatgagatat	1140
DB	1081	CTGTGAGCTTTGTAAGAAGTACTATTATTATTTATTTATTTGTGACAAATGTTAAATGATAT	1140
OY	1141	taaaagagaataaatcatgattctcttc	1168
DB	1141	TAAAGAGATAATATCATGATTCTCTTC	1168
RESULT 2			
AC069459/c		203083 bp	DNA linear HTG 27-JUN-2001
LOCUS	AC069459		
DEFINITION	Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT		
SEQUENCE	7 unordered pieces.		
ACCESSION	AC069459		
VERSION	AC069459.23		
KEYWORDS	GI:14547768		
SOURCE	HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULITOP.		
ORGANISM	house mouse.		
	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 203083)		
AUTHORS	Metzker,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C.,		
	Addison,S., Pace,A., Williams,C., Bonnin,D., Brooks,A., Garner,T.,		
	Buhay,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z.,		
	Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S.,		
	Fernandez,C., Ferraguto,D., Forcum-Tansey,J., Gill,R.,		
	Gorrell,J.H., Gunatane,P., Haller,G., Hernandez,J., Hogues,M.,		
	Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., Kelly,S.,		
	Kovar,C., Liu,J., Liu,W., Lonsheed,H., Lozano,R.J., Martin,R.,		
	Massey,E., McLeod,M.P., Mel,G., Moore,S., Morgan,M., Morris,S.,		
	Neel,D., Nelson,A., Nguyen,R., Nguyen,N., Ogun,M., Parish,B.,		
	Perez,L., Relter,D., Say,J., Shen,H., Vasquez,L., Watlington,S.,		
	Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A.,		
	Munhy,D.M., Rives,M., Scherer,S., Sodergren,E., Welstock,G.,		
	Worley,K. and Gibbs,R.		
TITLE	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 203083)		
AUTHORS	Worley,K.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-MAY-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
COMMENT	On Jun 25, 2001 this sequence version replaced gi:12621364.		
	----- Genome Center		
	Center: Baylor College of Medicine		
	Center code: BCM		
	Web site: http://www.hgsc.bcm.tmc.edu/		

Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: MAFO
 Center clone name: RP23-168P5
 Summary Statistics
 Sequencing vector: M13; L08821
 Chemistry: Dye-primer Bodyfly: 48% of reads
 Chemistry: Dye-terminator Big Dye: 52% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 212648 bases at least Q40
 Consensus quality: 218902 bases at least Q30
 Consensus quality: 222384 bases at least Q20
 Estimated insert size: 210556; sum-of-coverage estimation
 Quality coverage: 0x in Q20 bases; average-IP estimation
 Quality coverage: 7.2x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 62152: contig of 62152 bp in length
 * 62153 62252: gap of unknown length
 * 62253 118772: contig of 56520 bp in length
 * 118773 118872: gap of unknown length
 * 118873 148924: contig of 30052 bp in length
 * 148925 149024: gap of unknown length
 * 149025 167231: contig of 18207 bp in length
 * 167232 167332: gap of unknown length
 * 167332 189907: contig of 22576 bp in length
 * 189908 190007: gap of unknown length
 * 190008 196537: contig of 6530 bp in length
 * 196538 196637: gap of unknown length
 * 196638 203083: contig of 6446 bp in length.

FEATURES
 Source
 1. 203083
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-168P5"

BASE COUNT 52662 a 49293 c 47892 g 52633 t 603 others
 ORIGIN

Query Match 60.9%; Score 711; DB 2; Length 203083;
 Best Local Similarity 99.6%; Pred. No. 2,1e-175;
 Matches 744; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

OY 425 caggtgacttgatgaaggaagagctgtacctaagctgactgtgtgaacgt 484
 |||||||
 DB 43019 CAGGTGACTTGTATGAGGAAAGCGTCTACCTGAGCTGCTGGAACGT 42960
 |||||||
 OY 485 gtagctgacctgctgtctgaagaattctcagccagcagaagcttctctggccc 544
 |||||||
 DB 42959 GTCCTGGCCCTGGCTGCGCCGGAAGATTCACGACACACAGCAAGCTCTCTGGGCC 42900
 |||||||
 OY 545 cagctcgtttgtcagagtgatctggagctgtccgctggcagagtgcttcccttcg 604
 |||||||
 DB 42899 CAGCTCGTTTGTGCGACAGGTGTGGGCTTTCCCGCTGGCCAGGGTTTCCCTTGG 42840
 |||||||
 OY 605 atccgacccctccctgggtcactcttaagctgtcccccctcttaacctacttgactc 664
 |||||||
 DB 42839 ATCCGACCCCTCCCTGGGCTCATCTTAAGGCTGCCCTTCTTAACCTTTGGACTC 42780
 |||||||
 OY 665 ttccagctcactgaagggcctgtctcccaagatccctaacttccctgtctcag 724
 |||||||
 DB 42779 TTTCAAGTTTCACTAGAGGGGCTTGTCTCCCAAGATTCCTTAACCTTCCCTGGCTCCAG 42720

OY 725 agcatcaccaacctccctaccccccaacctccctccacccctcgtctgtcttgc 784
 |||||||
 DB 42719 AGCATCACCAACACTCCCTGCTACCCACCCACCTCTCCACCCCTGCTCTTGGTC 42660
 |||||||
 OY 785 cagctcgtt-cttctcctaaagcagcagagctgttcaatg-ttccatccacaga 842
 |||||||
 DB 42659 CAGCTCGTCTCTCTCTCAAGGACAGCAGAGCTGTTCACATGTTTTCATTCACAGCA 42600
 |||||||
 OY 843 cgtatctgtcttctt-cttaactcccatcccaacaataatccactactagctcc 901
 |||||||
 DB 42599 CGTATCTTGTCTCTCTTAACTCCATCCACCAAACTATCCACTTCACTAGCTCC 42540
 |||||||
 OY 902 aagccctactatccctcctaccccccacacacacacacacacacacacacac 961
 |||||||
 DB 42539 AAGCCCTCTATCTATCTCTGACCTCCCAACCACTACCGGACGACTGTTATGACTT 42480
 |||||||
 OY 962 tgtcaccaagcactgagatggcctgagcctgtgtgcaagaagcagaagacactg 1021
 |||||||
 DB 42479 TGTGACACAGGACACTGAGTGGCTGGACCTGGTGAGAGAACCAAGACCTGGGACT 42420
 |||||||
 OY 1022 aggcagaagcttcccaactgtgagggaggaagactgaggaacagctccctcgtatcc 1081
 |||||||
 DB 42419 AGGCAGAAAGTCCCACTGTGAGGGGAGAGACTGGGACAAAGCTCTCTCTGAGTCC 42360
 |||||||
 OY 1082 tgtgatttgaagaatactatttattattatgtgacaaatgttaagatatt 1141
 |||||||
 DB 42359 TGTGATTTTGAAGAAGATCTATTATTATATTGTGCAAAATGTTAAATGATATT 42300
 |||||||
 OY 1142 aagaagaataatcatgattctcttc 1168
 |||||||
 DB 42299 AAAGAGATTAATCATGATTTCTCTTC 42273

RESULT 3
 AL603707/c 234182 bp DNA linear ROD 17-NOV-2001
 LOCUS Mouse DNA sequence from clone RP23-422L16 on chromosome 11,
 DEFINITION complete sequence.
 ACCESSION AL603707
 VERSION AL603707.5 GI:17017790
 KEYWORDS house mouse.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 234182)
 REFERENCE
 AUTHORS Pearce, A.
 TITLE Direct Submission
 JOURNAL Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On Nov 20, 2001 this sequence version replaced gi:16603765.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30).
 An attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em., EMBL; Sw.,
 SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
 database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-422L16 is
 from the RPCR-23 Mouse PAC Library
 constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>

COMMENT

VECTOR: pBACE3.6
This sequence is the entire insert of clone RP23-422L16.
Location/Qualifiers
1. .234182
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_11p="RPC1-23"
complement(84050..84131)
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."

misc_feature
BASE COUNT 59310 a 56824 c 57519 g 60529 t
ORIGIN

Query Match 60.9%; Score 711; DB 10; Length 234182;
Best Local Similarity 99.6%; Pred. No. 2.1e-175;
Matches 744; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

425 caggtgacatttgatgaggaagagctgtctactgaagctgtagtgaacgt 484
66233 CAGGTGACATTGATGAGGAGAAAGCTGTCTACTGAAGCTGTGTAACGCT 66264
485 gtgctggcctcgctgcctggaagaaattctcaagccaaagcagcagctcctgagcc 544
66263 GTGCTGGCCCTCGCTGCCTGGAAGAAATTCTACGCCAACAGCAGAGCTCTCTGGGCC 66204
545 cagctccgtttgtgcagaggtgtctgggtgtgtgcgctgagcagaggtctcccttgg 604
66203 CAGCTCCGTTTGTGCAGAGGTGTGCGGTGTGCGCTGCGCCAGAGGTCTCTCCCTTGG 66144
605 atccgacaccccccctggcctgcatcttaaggctgccccctccaaaccttctggaactc 664
66143 ATCCGACACCCCTGGCCTGATTAAGGCTGCCCCCTTCTTAACCTACTTTGGAATC 66084
665 ttcaagttcaatgaggggctgtctctccagattctcaaatctccctgtgctcagag 724
66083 TTTCAAGTTCAATGAGGGGCTGTCTCTCCAGATTCTTAACCTTCCCTGGCTCCAGG 66024
725 agcatcaacacacccctccaccccccaacctctccacccccctgctgtctcttggc 784
66023 AGCATCAACACACCTCCCTACCCCAACCCCACTCTCCACCCCTGCTGCTCTGCTG 65964
785 cagctcgtt-ctctcctcaaaagcagcagagctgttcaatg-cttcaattccacaga 842
65963 CAGTCTGCTCTCTCTCAAAAGCAGCAGAGCTGTTCATGTTTTCATTCACAGCA 65904
843 cgtatcctgtctctc-ctaaatcccaatcccaacaaactatccactcaatgctccg 901
65903 CGTATCTCTCTCTCTTAAATCCATCCATCCACACAACTATCCACTCTCACTAGCTCC 65844
902 aaagccctacttaccctgactcccccacacactcaccagcagcagctgttatgactt 961
65843 AAAGCCCTACTTATCTCTGACTCCCAACCACTCACTCACTCACTCACTCACTCACT 65784
962 tgttcacacagcagctgagatgagctgtgacctgtgtgcagagacagagaaactcggact 1021
65783 TGTTCACACAGCAGCTGAGATGAGCTGTGACCTGTGTGCGAGAGCAAGAACTGCGACT 65724
1022 aggcacaagatgccaaactgtgagggggaagagctgggagacaagctccctcgtgacc 1081
65723 AGGCACAAGATGCCAAACTGTGAGGGGGAAGAGCTGGGGAACAAGCTCTCCCTGGATCC 65664
1082 tgttgatttgaagaaatactatttatttatttgaacaaatgttaattgatatc 1141
65663 TGTGATTTTGAAAGAAATATTTTATTTATTTATTTGACAAATGTTAAATGATATT 65604
1142 aaagagaataaatcatgattctctc 1168
65603 AAAGAGAATAAATCATGATTCTCTC 65577

RESULT 4
AX201324 1353 bp DNA linear PAT 30-AUG-2001
LOCUS
DEFINITION
Accession
AX201324
VERSION
AX201324.1 GI:15391154
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1353)
Ashkenazi, A.J., Goodard, A., Godowski, P.J., Gurney, A.L.,
Hillan, K.J., Masters, S.A., Pan, J., Pitt, R.M., Roy, M.A., Smith, V.,
Stone, D.M., Watanabe, C.K. and Wood, W.I.
Compositions and methods for the treatment of tumour
Patent: WO 0153486-A 3 26-JUL-2001;
Genentech, Inc. (US)
TITLE
JOURNAL
FEATURES
source
1. .1353
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 257 a 443 c 389 g 264 t
ORIGIN

Query Match 53.8%; Score 628.6; DB 6; Length 1353;
Best Local Similarity 76.2%; Pred. No. 4.7e-154;
Matches 933; Conservative 0; Mismatches 219; Indels 73; Gaps 10;

2 gtgtgagcctggcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 61
130 GCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 189
62 ggagagcctggagagcctgctgctgctgctgctgctgctgctgctgctgctgctgctg 121
190 GGGAGCGGAGATGCTGCTCGCCAGAGGCTCCAGAGGAGCTGGGCTGGGCTGGGCTGG 249
122 cgcggagagcccttgactgaactgaatcccaagacagaagaaacagagctgtgtaccttc 181
250 GACGAGAGCCGCTGGAACTGAATCCCAAGACAAAGAACCCAGATCCCTGGCTTTC 309
182 ttgaaacaactagtcgagcctcgaagaagtgctcctaaagccggaagcgcgctcgc 241
310 CTGAACGACATGATGCGGCTCGCAGAAATGACACTTAAAGGCCGGAACACGGGCTCGA 369
242 cgaactatgagcccaattagatgactatcctggcagagcagagctgagagcagaaga 301
370 AGACGATCGGACCCATTAATGAAATTCATCAGACCTGACAGAGAGGAGCGCAGGCA 429
302 ggtgtgagctgagagctgagctgagctgaggaagagacaaatcaacacagctccagcctctg 361
430 GGTGTGAGCGGAGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 489
362 cgtacgacgcagatgtgggaattcaagtcacagagctgagctgctactactgttac 421
490 CGCTAACACCCCGAGATCGGAGATTATAGTCAACCGGCTGCGGCTACTACTCTGTAC 549
422 tgtcaagagcacttgatgaggaagagctgtctcaacgaagcagagctgtgtgtgaaac 481
550 TGTCAAGGCACTTGTGATGAGAGGAGCTGTCTTACCTGAAGGCTGTGCTGTGTGAT 609
482 ggtgtgagcctggcctgctgctggaagaattctcagccacagcagcagagctccctggg 541
610 GGTGTGAGCGGCTGCTGCTGCTGAGAGAAATTCACAGCACTGGGCGAAGTTCCTGGG 669
542 cccagagctcgtgtgtgcaagtgctgtgagctgtgtgctgctgctgctgctgctgctcct 601
670 CCCAGAGCTCCGCTGTGCAAGTGTGAGCTGTGTGAGCTGTGTGAGCTGTGTGAGCTGT 729
602 cggatccgacacccctcctgagctcatttcaagctgagcctccctcctcaactacttggga 661
730 CGATCCGACACCTTCCCTGGGCTGCTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789

Db	Accession	Gene	Protein	Length	Start	End	Score	E-value
Db	850	-----CTGCGAAGCTCTCTGTGGGACCCGGTCCCTCTGGCCACCCGACGGCGTCTTT		904				
Qy	781	ggtccagtcctgctctctcc--tcaaggcagccagagctgttcaagyltccatcc-		837				
Db	905	GCTCGAGACCTGCCCCCTCCCTCTTGAAGCGCTGCGGCGCTTTCACAGTTCATCC		964				
Qy	838	-----aagaaagatctctgtctctcttcaatccatcccaactatccatcc		891				
Db	965	ACATTAATTAAGATTAATCCCACTCTTAATCTTAACACTCCCCACGCGCACTCTCACCTC		1024				
Qy	892	actagctcccaagccctactat-----ccctgactcccccaccat		936				
Db	1025	ACTAGCTCCCCCAATCCCTGACCCCTTTGAAGCCCCCAGTATCTGCAGCTCCCCCTGGCCA		1084				
Qy	937	caaccgaccagctgttattgaactttgtgac-----		968				
Db	1085	CAGACC-----CAGAGCTCATTTGGTTCACTGTACTGTGGCAAGATGGTCCAGAACCC		1144				
Qy	965	-----caggcaactgagatgagctggaactgtgtgaggaagccagaagactgtgactag		1023				
Db	1145	CAGTTCAGGCACTAAGAGGGGCTGAGACCTGGCGGAGGAAAGACAGACTGGGCTTAG		1204				
Qy	1024	gccagaagctcccaactgtggaagggaagagactggaagccctccctgtga-----tc		1079				
Db	1205	GCCAGAGAGTTCCTCCCAATGTGAGGGCGGAGA-AACAGACAACTCTCTTGTGAGAAATTC		1263				
Qy	1080	cctgtggaatttgaagaa--agatactatttatttatttatttgaacaaatgt--taaat		1134				
Db	1264	CCTGTGAGATTTTAAACAGATATTATTATTATTATTATTGTGACAAATATTGATTAAT		1323				
Qy	1135	ggatattaaagagaataatcatga		1159				
Db	1324	GGATATTAAATGAAATAGTCACTATA		1348				

RESULT 6

AF030099 LOCUS Homo sapiens TWEAK mRNA, complete cds. linear PRI 20-DEC-1997

AF030099 DEFINITION Homo sapiens TWEAK mRNA, complete cds.

AF030099 ACCESSION AF030099

AF030099 VERSION AF030099.1 GI:2707218

KEYWORDS human.

SOURCE ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1306) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 2 (bases 1 to 1306) Chicheportriche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H., Hession,C., Garcia,I. and Browning,J.L. TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis J. Biol. Chem. 272 (51), 32401-32410 (1997) 98070415

JOURNAL 2 (bases 1 to 1306) Bourdon,P., Hession,C., Tizard,R. and Browning,J. Direct Submission Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA

FEATURES

source Location/Qualifiers

1..1306 /organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="17"

/map="17p13"

/tissue_type="tonsil"

18..767 /note="Ligand in the TNF family; secreted protein; start codon not verified experimentally"

/product="TWEAK"

/codon_start=1

/protein_id="AAC51923.1"

/db_xref="GI:2707219"

/translation="MAARSRQRGRGRGPGSTALVPLALGLALACLGILLAVSIL"

[illegible]


```

Oy 418 gtaetgtcaggtgacatttgatgaggaaagctgttactgaagctgagctgtgct 477
Db 833 ttttttttttttttttttttttttttttttttttttttttttttttttttttttt
Oy 478 gaacggtgtctgtgctgtgctgtgctgtgctgtgctgtgctgtgctgtgct 537
Db 893 ggatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
Oy 538 tgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 597
Db 953 cgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1012
Oy 598 cctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 657
Db 1013 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1072
Oy 658 tgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 717
Db 1073 cgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1132
Oy 718 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 776
Db 1133 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1187
Oy 777 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 834
Db 1188 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1247
Oy 835 tttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 887
Db 1248 tttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1307
Oy 888 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 932
Db 1308 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1367
Oy 933 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 968
Db 1368 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1427
Oy 969 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1019
Db 1428 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1487
Oy 1020 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1077
Db 1488 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1546
Oy 1078 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1130
Db 1547 cttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1606
Oy 1131 aaatgatatttaataaataaataaataaataaataaataaataaataaataa 1159
Db 1607 aaatgatatttaataaataaataaataaataaataaataaataaataaataa 1635

RESULT 10
AC016876/c 190358 bp DNA linear HTG 03-SEP-2001
LOCUS Homo sapiens clone RP11-186B7, WORKING DRAFT SEQUENCE, 16 unordered
pieces.
AC016876 AC016876 GI:15421989
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.
REFERENCE
1 (bases 1 to 190358)
Bairren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Brown,A., Castle,A., Collins,S., Collins,S., Collymore,A.,
Cooke,P., Dearlano,K., Dewar,K., Domini,M., Donelan,L., Doyle,M.,
Galligan,J., Gardys,S., Grant,G., Hagos,B., Heath,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,B., Karlas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliou,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE
JOURNAL
COMMENT

```

```

2 (bases 1 to 190358)
Bairren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Brown,A., Castle,A., Collins,S., Collins,S., Collymore,A.,
Cooke,P., Dearlano,K., Dewar,K., Domini,M., Donelan,L., Doyle,M.,
Galligan,J., Gardys,S., Grant,G., Hagos,B., Heath,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,B., Karlas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vassiliou,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE
JOURNAL
COMMENT
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 3, 2001 this sequence version replaced gi:13431059.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 13849
Center clone name: 186_B_7
----- Summary Statistics
Sequencing vector: M13; M7815; 3% of reads
Sequencing vector: Plasmid; n/a; 97% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 181928 bases at least Q40
Consensus quality: 185229 bases at least Q30
Consensus quality: 187335 bases at least Q20
Insert size: 168000; agarose-fp
Insert size: 188858; sum-of-coverage
Quality coverage: 8.1 in Q20 bases; agarose-fp
Quality coverage: 7.2 in Q20 bases.
NOTE: This is a working draft sequence. It currently
consists of 16 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 20703: contig of 20703 bp in length
20704 20803: gap of 100 bp
20804 21935: contig of 1132 bp in length
21936 22035: gap of 100 bp
22036 24202: contig of 2167 bp in length
24203 24302: gap of 100 bp
24303 28199: contig of 3897 bp in length
28200 28299: gap of 100 bp
28300 35021: contig of 6722 bp in length
35022 35121: gap of 100 bp
35122 40745: contig of 5624 bp in length
40746 40845: gap of 100 bp
40846 47136: contig of 6291 bp in length
47137 47236: gap of 100 bp
47237 58980: contig of 11744 bp in length
58981 59080: gap of 100 bp
59081 68164: contig of 9084 bp in length
68165 68264: gap of 100 bp
68265 94625: contig of 26361 bp in length
94626 94725: gap of 100 bp
94726 102564: contig of 7839 bp in length
102565 102664: gap of 100 bp
102665 116764: contig of 14100 bp in length

```

```

* 116765 116864: gap of 100 bp
* 116865 136408: contig of 19544 bp in length
* 136409 136508: gap of 100 bp
* 136509 156380: contig of 19872 bp in length
* 156381 156480: gap of 100 bp
* 156481 180631: contig of 24151 bp in length
* 180632 180731: gap of 100 bp
* 180732 190358: contig of 9627 bp in length.
Location/Qualifiers
source
1. 190358
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-186B7"
/clone_11b="RP11-11 Human Male BAC"
1. 20703
/note="assembly_fragment"
clone_end:5p6
vector_side:left"
misc_feature
20804..21935
/note="assembly_fragment"
22036..24202
/note="assembly_fragment"
24303..28199
/note="assembly_fragment"
28300..35021
/note="assembly_fragment"
35122..40745
/note="assembly_fragment"
40846..47136
/note="assembly_fragment"
47237..58980
/note="assembly_fragment"
59081..68164
/note="assembly_fragment"
68265..94625
/note="assembly_fragment"
94726..102564
/note="assembly_fragment"
102665..116764
/note="assembly_fragment"
116865..136408
/note="assembly_fragment"
136509..156380
/note="assembly_fragment"
156481..180631
/note="assembly_fragment"
180732..190358
/note="assembly_fragment"
clone_end:7f7
vector_side:right"
BASE COUNT 45173 a 48605 c 50635 g 44441 t 1504 others
ORIGIN
Query Match 26.0%; Score 303.8; DB 2; Length 190358;
Best Local Similarity 71.0%; Pred. No. 1.7e-68;
Matches 574; Conservative 0; Mismatches 162; Indels 73; Gaps 10;
QY 425 caggtcactttagatgaggaaaggtctctaccctgaagctgactgctgtgaacggt 484
DB 83053 CAGGTCTACTTTGATGATGAGGGAAGGCTGCTCTACTCTAAGCTGACTGTGGATGCT 82994
QY 485 gtgtgcgcctgcgtcgtcgtgaagaattctcagcacagcagcagctctcctggccc 544
DB 82993 GTGCTGCGCCTGGCTGCTGCGGGAATTCCTCAGCACATGCGGCGAGTTCCCTCGGCCC 82934
QY 545 cagctcgtttgtgcaggtgtcgtcgtgtgcgcgtcgcgcaggtcttcctcctgg 604
DB 82933 CAGCTCCGCTCTGCGCAGGTGCTGCTGCTGCGGCGGCGGCGGAGGCTCTCCCTGCGG 82874
QY 605 atccgacacccctccctgggctctcttaagctgccccctcttaacacttaacac 664
DB 82873 ATCCGACACCTCCCTGGCGCCATCTCAAGGCTGCCCTTCTCTCACTTCTGGAGCTC 82814

```

```

QY 665 tttaagttcactgaggggctgtctctccagatctcttaactctccctgtccagg 724
DB 82813 TTTCAGGTTCTACTGAGGGGCGCTGGTCTCCCGCAGTGTCTCCAGGTCGGGCTCC--- 82755
QY 725 agcatcaccacactctccctaccaccccaactctccctcaccctcc-ctgtctcctgt 783
DB 82756 --CTTCGACAGGCTCTCTGGGCAACCGGTCTCCCTGCGCCACCTCCAGCGGCTCTTGTCT 82699
QY 784 ccagtcctgtctctcc--ccaaggcagccagagctgtctacatgcttccatcc--- 837
DB 82698 CCAGACCTTCCCTCCCTCCCTTAAGAGCTGCTGCGCTGTTCACGTGTTTTCATCCACA 82639
QY 838 ---acaagcgtatctctgtctctcttaacatccatccacacacacttaacactcaact 894
DB 82638 TAAATACATATTCCACACTTATGCTTACACTCCGCCACCCGCCACATCTCACCCTCACT 82579
QY 895 agctcccaagccctac-----tatacctgtactcccccacactaac 939
DB 82578 ACTCTCCCAATCCCTGACCCCTTTGAGGCCCCCAAGTATCTGACATCTCCCTGCGCACAG 82519
QY 940 ccgaccagtgcttattgtactgtgac----- 968
DB 82518 ACCCCAGGGCATTTGTCTTCACTGTACTCTGTGGCAGATGGGTCCAGAAAGACCCAC 82459
QY 969 --caggcactgagatggtcgtgacctgtgycaggaagccagaagacctgtgactaagcc 1026
DB 82458 TTCAGGACATAAGAGGGGCTGTGACCTGTGGCGGAGCAAGACACTGGGCTTAGGGC 82399
QY 1027 aagaattcccaactgtgagggggaagagctgaggacaaagctctccctgga-----tcct 1082
DB 82398 AGAGTTCCTCCCAATGTGAGGGGCGAG--AACAAAGCAACACTCTCCCTTGAAATATTCCT 82340
QY 1083 gtggatttgaagaa--agatacatttatttattatctgtgacaaaatgt---taaatga 1137
DB 82339 GTGGATTTTAAACAGATATTATTTTATTTATTTATGTGACAAATGTTGATTAATGCA 82280
QY 1138 tattaagagaataaatacatgattctct 1166
DB 82279 TATTAAATAGAAATAGTATAGTCTCTCT 82251

RESULT 11
AC098923 179030 bp DNA linear HTG 20-DEC-2001
LOCUS AC098923/C
DEFINITION Rattus norvegicus clone CH230-154B15, *** SEQUENCING IN PROGRESS
ACCESSION AC098923
VERSION AC098923.4 GI:17973852
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 179030)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbarta,J.,
Benton,J., Blinage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Briveau,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,
Burich,P., Burrell,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthalite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Barnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frentz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Huijck,S., Hume,J.,

```

Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J., Kovar, C., Kratovic, J., Kureshi, R., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louieged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabhat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Ntkenko, S., Ogun, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rives, R., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabori, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 179030)

Morley, K.C.

Direct Submission

Submitted (06-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:17064677.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: G10K

Center clone name: CH230-154B15

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to findPhrapList

Consensus quality: 137096 bases at least Q40

Consensus quality: 145080 bases at least Q30

Consensus quality: 152950 bases at least Q20

Estimated insert size: 138980; sum-of-coverage estimation

Quality coverage: 0x in Q20 bases; agarose-IP estimation

Quality coverage: 2.2x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length

* (see <http://www.hgsc.bcm.tmc.edu/docs/genbank/draft-data.html>).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 68 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 7467: contig of 7467 bp in length

* 7468 7567: gap of unknown length

* 7568 12115: contig of 4548 bp in length

* 12116 12215: gap of unknown length

* 12216 20313: contig of 8098 bp in length

* 20314 20413: gap of unknown length

* 20414 25589: contig of 5176 bp in length

* 25590 25689: gap of unknown length

* 25690 31434: contig of 5745 bp in length

* 31435 31534: gap of unknown length

* 31535 37595: contig of 6061 bp in length

* 37596 37695: gap of unknown length

* 37696 42172: contig of 4477 bp in length

* 42173 42272: gap of unknown length

42273 46234: contig of 3962 bp in length

* 46235 46334: gap of unknown length

* 46335 50605: contig of 4271 bp in length

* 50606 50705: gap of unknown length

* 50706 55202: contig of 4497 bp in length

* 55203 55302: gap of unknown length

* 55303 58407: contig of 3105 bp in length

* 58408 58507: gap of unknown length

* 58508 61287: contig of 2779 bp in length

* 61288 61386: gap of unknown length

* 61387 64467: contig of 3081 bp in length

* 64468 64567: gap of unknown length

* 64568 68498: contig of 3931 bp in length

* 68499 68598: gap of unknown length

* 68599 73449: contig of 4851 bp in length

* 73450 73549: gap of unknown length

* 73550 77403: contig of 3853 bp in length

* 77403 77503: gap of unknown length

* 77503 81027: contig of 3525 bp in length

* 81028 81127: gap of unknown length

* 81128 84144: contig of 3017 bp in length

* 84145 84245: gap of unknown length

* 84245 86680: contig of 2436 bp in length

* 86681 86780: gap of unknown length

* 86781 89210: contig of 2430 bp in length

* 89211 89310: gap of unknown length

* 89311 92339: contig of 3029 bp in length

* 92340 92439: gap of unknown length

* 92440 94951: contig of 2512 bp in length

* 94951 95051: gap of unknown length

* 95052 95052: contig of 3535 bp in length

* 95053 98587: gap of unknown length

* 98587 98686: gap of unknown length

* 98687 100597: contig of 1911 bp in length

* 100598 100697: gap of unknown length

* 100698 102766: contig of 2069 bp in length

* 102767 102866: gap of unknown length

* 102867 105237: contig of 2371 bp in length

* 105238 105337: gap of unknown length

* 105338 107305: contig of 1968 bp in length

* 107306 107405: gap of unknown length

* 107406 109057: contig of 1652 bp in length

* 109058 109157: gap of unknown length

* 109158 111454: contig of 2297 bp in length

* 111455 111554: gap of unknown length

* 111555 113704: contig of 2150 bp in length

* 113705 113804: gap of unknown length

* 113805 115933: contig of 2129 bp in length

* 115934 116033: gap of unknown length

* 116034 118193: contig of 2160 bp in length

* 118194 118293: gap of unknown length

* 118294 120414: contig of 2121 bp in length

* 120415 120514: gap of unknown length

* 120515 122311: contig of 1797 bp in length

* 122312 122411: gap of unknown length

* 122412 125628: contig of 3217 bp in length

* 125629 125728: gap of unknown length

* 125729 128545: contig of 2817 bp in length

* 128546 128645: gap of unknown length

* 128646 130563: contig of 1918 bp in length

* 130564 130663: gap of unknown length

* 130664 132945: contig of 2282 bp in length

* 132946 133045: gap of unknown length

* 133046 134556: contig of 1511 bp in length

* 134557 134656: gap of unknown length

* 134657 136478: contig of 1822 bp in length

* 136479 136578: gap of unknown length

* 136579 138159: contig of 1581 bp in length

* 138160 138259: gap of unknown length

* 138260 140456: contig of 2197 bp in length

* 140457 140556: gap of unknown length

* 140557 142544: contig of 1988 bp in length

* 142545 142644: gap of unknown length

* 142645 143784: contig of 1140 bp in length

Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogun,M., Okwunonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshbari,N.,
Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,F., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,D., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vanson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 118395)
Worley,K.C.

Direct Submission
Submitted (07-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:16756150.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GIJO
Center clone name: CH230-191M14
----- Summary Statistics -----
Assembly program: Phrap; version 0.990329First call to
findphraplist
Consensus quality: 85618 bases at least Q40
Consensus quality: 94111 bases at least Q30
Consensus quality: 99937 bases at least Q20
Estimated insert size: 55807; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-1p estimation
Quality coverage: 0.7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 68 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	3372: contig of 3372 bp in length	28086	29145: contig of 1060 bp in length
3373	3472: gap of unknown length	29146	29245: gap of unknown length
3473	7021: contig of 3549 bp in length	29246	31057: contig of 1812 bp in length
7022	7121: gap of unknown length	31058	31157: gap of unknown length
7122	9925: contig of 2804 bp in length	31158	33961: contig of 2804 bp in length
9926	10025: gap of unknown length	33962	34061: gap of unknown length
10026	12351: contig of 2326 bp in length	34062	35749: contig of 1688 bp in length
12352	12451: gap of unknown length	35750	35849: gap of unknown length
12452	15703: contig of 3252 bp in length	35850	37495: contig of 1646 bp in length
15704	15803: gap of unknown length	37496	37595: gap of unknown length
15804	18388: contig of 2585 bp in length	37596	39453: contig of 1858 bp in length
18389	18488: gap of unknown length	39454	39553: gap of unknown length
18489	21644: contig of 3156 bp in length	39554	41475: contig of 1922 bp in length
21645	21744: gap of unknown length	41476	41575: gap of unknown length
21745	24057: contig of 2313 bp in length	41576	43934: contig of 2359 bp in length
24058	24157: gap of unknown length	43935	44034: gap of unknown length
24159	25917: contig of 1760 bp in length	44035	45739: contig of 1705 bp in length
25918	26017: gap of unknown length	45740	45839: gap of unknown length
26018	27985: contig of 1968 bp in length	45840	47091: contig of 1252 bp in length
27986	28085: gap of unknown length	47092	47191: gap of unknown length
		47192	48852: contig of 1661 bp in length
		48853	48952: gap of unknown length
		48953	50055: contig of 1103 bp in length
		50056	50155: gap of unknown length
		50156	51577: contig of 1422 bp in length
		51578	51677: gap of unknown length
		51678	53662: contig of 1995 bp in length
		53663	53762: gap of unknown length
		53763	55525: contig of 1763 bp in length
		55526	55625: gap of unknown length
		55626	56640: contig of 1015 bp in length
		56641	56740: gap of unknown length
		56741	58876: contig of 2136 bp in length
		58877	58976: gap of unknown length
		58977	60443: contig of 1467 bp in length
		60444	60543: gap of unknown length
		60544	61942: contig of 1399 bp in length
		61943	62042: gap of unknown length
		62043	63377: contig of 1335 bp in length
		63378	63477: gap of unknown length
		63478	64892: contig of 1415 bp in length
		64893	64992: gap of unknown length
		64993	67304: contig of 2312 bp in length
		67305	67404: gap of unknown length
		67405	68639: contig of 1235 bp in length
		68640	68739: gap of unknown length
		68740	69861: contig of 1122 bp in length
		69862	69961: gap of unknown length
		69962	71276: contig of 1315 bp in length
		71277	71376: gap of unknown length
		71377	72603: contig of 1227 bp in length
		72604	72703: gap of unknown length
		72704	74279: contig of 1576 bp in length
		74280	74379: gap of unknown length
		74380	76194: contig of 1815 bp in length
		76195	76294: gap of unknown length
		76295	77477: contig of 1183 bp in length
		77478	77578: gap of unknown length
		77579	78686: contig of 1109 bp in length
		78687	78786: gap of unknown length
		78787	80295: contig of 1509 bp in length
		80296	80395: gap of unknown length
		80396	81556: contig of 1161 bp in length
		81557	81656: gap of unknown length
		81657	82923: contig of 1267 bp in length
		82924	83023: gap of unknown length
		83024	84800: contig of 1777 bp in length
		84801	84900: gap of unknown length
		84901	86054: contig of 1154 bp in length
		86055	86154: gap of unknown length
		86155	88128: contig of 1974 bp in length
		88129	88228: gap of unknown length
		88229	89419: contig of 1191 bp in length

*	88420	88519:	gap of unknown length
*	88920	91020:	contlg of 1501 bp in length
*	91021	91120:	gap of unknown length
*	91121	92252:	contlg of 1132 bp in length
*	92253	92352:	gap of unknown length
*	92353	94449:	contlg of 2097 bp in length
*	94450	94549:	gap of unknown length
*	94550	95619:	contlg of 1070 bp in length
*	95620	95719:	gap of unknown length
*	95720	97175:	contlg of 1456 bp in length
*	97176	97275:	gap of unknown length
*	97276	98304:	contlg of 1029 bp in length
*	98305	98404:	gap of unknown length
*	98405	99763:	contlg of 1359 bp in length
*	99764	99863:	gap of unknown length
*	99864	101441:	contlg of 1578 bp in length
*	101442	101541:	gap of unknown length
*	101542	102713:	contlg of 1172 bp in length
*	102714	102813:	gap of unknown length

Query Match	4.88;	Score 56.2;	DB 2;	Length 118395;
Best Local Similarity	48.98;	Pred. No. 0.0017;		
Matches 151; Conservative	0;	Mismatches 158;	Indels 0;	Gaps 0;

Qy	Db	Qy	Db
639	cccccttccaaacacttacttgaaactcttcaagttaacttggagggtcttgcctccaga	639	cccccttccaaacacttacttgaaactcttcaagttaacttggagggtcttgcctccaga
68286	ccnctttgctcccccccccccccccccccttccccctctctcttggctatctcccccc	68286	ccnctttgctcccccccccccccccccccttccccctctctcttggctatctcccccc
699	ttcccttaaacattccctgtgctccagagatcaacacactctccttaacccaccccaactc	699	ttcccttaaacattccctgtgctccagagatcaacacactctccttaacccaccccaactc
86346	ttcccttccccctgttcccccccccccccccccttcccccttccccctcccccccccccccc	86346	ttcccttccccctgttcccccccccccccccccccccttcccccttccccctcccccccccccccc

Dy 759 ctcacccccctcgcgtcccttgctccagttccttctctccccaaggacgaccagact 818
| | | | | | | | | | | | | | | |
Db 86406 ccccccctcccccctccmcccttcccccctcccmcccccctcccccctatccmc 86465

Qy 819 gtltcaatgtttccattccacagacgatatccttgccttccttaacataccatcccaccac 878
+ + + + + + + + + + + + + + + +
Db 86466 TCCCCCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCTCCCCCCCCCCCCCCCCCCC 86523

Dy 879 aaatatacaacactaacctagctccccaagaacctattatcctcgtagtccccaccactca 938
| | | | |
Db 86526 ccccccccccccccccctccccccccccccccctccccccccccccctccccac 86585

QY 939 ccgcagccac 94/
||| || |
Db 86586 ccctccctc 86594

| LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS |
|---------------------|---|-----------|--------------------|------------------------|
| SULT_14
AF429315 | Homo sapiens junctionophillin 3 (JPH3) gene, partial cds. | AF429315 | AF429315.1 | GI:17646244 |
| LOCUS | AF429315 | 125020 bp | DNA | linear PRI 18-JAN-2002 |
| DEFINITION | Homo sapiens junctionophillin 3 (JPH3) | | gene, partial cds. | |
| ACCESSION | AF429315 | | | |
| VERSION | AF429315.1 | | | GI:17646244 |
| KEYWORDS | | | | |
| SOURCE | human. | | | |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 125020)

TITLE Ingersoll-Ashworth, R.G., Fieisher, A., Stevanin, G., Brice, Potter, N.T., Ross, C.A. and Margolis, R.L.
A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington's disease

| | |
|-----------|------------------------------------|
| JOURNAL | Nat. Genet. 29 (4), 377-378 (2001) |
| MEDLINE | 21583737 |
| PUBMED | 11694876 |
| REFERENCE | 2 / bases 1 to 125020) |

AUTHORS Holmes, S.E., Ingersoll-Ashworth, R.G., Ross, C.A. and Margolis, R.L.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical

```

Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
location/Qualifiers
1. 125020
/coordinates="None detected"

```

```

/map="16q24.3; between D16S520 and WI-12410"
/note="Isolated from a patient with Huntington's
disease-Like 2 (HDL2)"
complement(35581..35746)
repeat_region
/next twooctadom

```

| repeat_region | complement(35581..35746) |
|---------------|--|
| mRNA | /rpt_type=tandem
/rpt_unit=ctg
complement(<36507..>36687)
/ctg="TTCG" |

| | |
|------|--|
| mRNA | complement(<36507. .>36887)
/gene="JPH3" |
| gene | /product="junctophilin 3"
complement(<36507. .>36887)
/gene="JPH3" |

```
gene      complement(<36507. .>36887)
          /gene="JP_H3"
          /note="JP3"
          complement(<36507. .36887)
CDS       /locus="JP_H3"
```

| Gene | Protein | Function |
|------|--|--|
| CDS | complement(<36507. .36887)
/gene="JPH3" | component of the junctional complex between plasma membrane and endoplasmic reticulum" |

```

/codon_start=1
/product="juncatophilin 3"
/protein_id="AAL40941.1"
/db_xref="GI:17646245"

```

```

/db_xref="GI:17646245"
/translation="MSGGRFNPDDGSGYCGGMEGRKAHNGVCTGKQGGETGSG
HGFEVLVYMPGSGNTYOGTMAQGRKHGIGLESKRWYKKEWTHGKRGKRGVEEC
NGAYEECTMSGNGDQCGTGYTSDG"
20056..20731..20802 + 1254..20799

```

| | a | c | g | t | others |
|------------|-------|-------|-------|-------|--------|
| BASE COUNT | 29056 | 32731 | 30696 | 28283 | 4254 |
| ORIGIN | | | | | |

| | | | | |
|-----------------------|-------------------|-------------------|-----------|----------------|
| Query Match | 4.8%: | Score 56.2: | DB 9: | Length 125020: |
| Best Local Similarity | 14.4%: | Pred. No. 0.0017: | | |
| Match 82: | Conservative 328: | Mismatches 240: | Totals 8: | Gaps 2: |

Qy 519 ccacgcagcgaagctctcctcgtgccccacgctccggtttgtccaggtgtcttgagctgttc 578

DQ 579 cgcctcgccagcgggtcttccttcctcgatccgacccctccctgygcctcatcttaagctg 633
:::| |::::: || :| : : : | | : | | : : | : : |
Db 16968 MRYTGSKMMSSMCTSSCYAMCMCCMCWSCCNRSCCCNCRSYCCMRYYCA--CKG 170

QY 639 ccccttcctaacctactttggaactcttccagttccactgagggccttgcctctccaga 6398
 :::::|: : : :::::| : : :|:::| : : : :
 Db 17025 YMSSTWTMSAYSRRLYWKRSKMKCWSRMCSSRSRSKCKSRGGGSGNKGKGSYCKRT 17025

```
Qy 699 ttccctaactttccctgyc-----tccagagcatcaccaacttcctaccaccacc 75
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17085 KRKSRRGMKGAQWYMYRRSRMRKMKMYSSGMYCNCYCMQWRGRCYCSCMTSSAMCSY 17
```

Oy 754 cactcctccacccccctgcgtccttcttcagtcctgtctctctccaagaagcagcag 81
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : ||
Db 17145 CYAKCSMCYSCVYGSMKGGYYIACSYRSMSSKKCMRGSTYSTGCCCTTTTTCCCC 172

Qy 814 agctggttcaacatgtttccattccacagacgtatcctgtcctctcttaacalccatccc 87
| | | | | : : : | | : : | : | : :
Db 17205 NANTGGGGAAGCTTTTGNKKTYSYRKRRNGCAMKYNNYNNSWRSSCRAGMSCTKYKS 172

Qy 874 accacaataatccactcaactagctcccagaagccctactatccccgacccccacc 93
 : ::::: - : - : - : : : : : : - : - : - : - :
Db 17265 SMTMSASYCWCMSMYCSCMRMASRGWSSYMKMMKSSPRMSTYKKCCWCCKMCYCMM 17

Qy 934 actcaccgcagacagctgttatgacttctgtcaccaaggcactgatatggcgagacctg 99
:: :: :: :: : : : : : : : : |
Db 17325 RSMRSGMSYIMYASWSSSRGCTRCYWCMSKSKCYKSYMMMRBKRMGKMKWSRCWG 173

QY 994 gtgcaaggaagcagagaacatctgactgcagcaagtcccaactgtgaggggaga 10
 :|::: : ::|: |: |::|::|::|::|:
Db 17385 SMMGASRRSSCTCYKYSMRCSMNSSKCYCAGCMMKGGYMRYMRCWSMKRRRWGKSANY 174

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2002, 11:56:14 ; Search time 3297.36 Seconds
(Without alignments)
8713.678 Million cell updates/sec

Title: US-09-245-198a-3
Perfect score: 1373
Sequence: 1 atgcatgttgtagcttga.....gacaaatgtgataatg 1373

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Genemdb1:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Length | DB ID | Description |
|------------|-------------|--------|-------|-------------|
|------------|-------------|--------|-------|-------------|

| | | | | | |
|----|--------|------|--------|----|----------|
| 1 | 1320.2 | 96.2 | 1353 | 6 | AX201324 |
| 2 | 1320.2 | 96.2 | 1368 | 9 | AF055872 |
| 3 | 1285 | 93.6 | 1306 | 9 | AF030099 |
| 4 | 1226.4 | 89.3 | 1236 | 6 | AR140407 |
| 5 | 1096.8 | 79.9 | 1651 | 9 | BC019047 |
| 6 | 768.2 | 56.0 | 190358 | 2 | AC016876 |
| 7 | 629.2 | 44.8 | 898 | 6 | AX180714 |
| 8 | 614.6 | 45.8 | 1168 | 10 | AF030100 |
| 9 | 309.6 | 22.5 | 203083 | 2 | AC069459 |
| 10 | 309.6 | 22.5 | 234182 | 10 | AL603707 |
| 11 | 158.6 | 11.6 | 179030 | 2 | AC098923 |
| 12 | 87.2 | 6.4 | 7218 | 6 | 166494 |
| 13 | 77.6 | 5.7 | 125020 | 9 | AF429315 |
| 14 | 74.8 | 5.4 | 303091 | 2 | AC084799 |
| 15 | 70.8 | 5.2 | 27571 | 2 | AC105667 |
| 16 | 70.8 | 5.2 | 73948 | 2 | AC022556 |
| 17 | 70.6 | 5.1 | 144979 | 2 | AC016280 |
| 18 | 70.4 | 5.1 | 220469 | 2 | AC074307 |
| 19 | 70.2 | 5.1 | 298166 | 2 | AC087563 |
| 20 | 69.2 | 5.0 | 936 | 8 | CNS01AYK |
| 21 | 69.2 | 5.0 | 68631 | 2 | AC094811 |
| 22 | 69.2 | 5.0 | 82897 | 2 | AC022595 |
| 23 | 68.2 | 5.0 | 110737 | 2 | AC011105 |
| 24 | 68.2 | 5.0 | 265537 | 2 | AC087228 |
| 25 | 68 | 5.0 | 300695 | 2 | AC079431 |
| 26 | 67.6 | 4.9 | 626 | 8 | CNS018RP |
| 27 | 67.2 | 4.9 | 66624 | 2 | AC101520 |
| 28 | 67.2 | 4.9 | 147124 | 2 | AC022768 |
| 29 | 67.2 | 4.9 | 177883 | 2 | AC046159 |
| 30 | 67.2 | 4.9 | 181988 | 2 | AC090552 |
| 31 | 67 | 4.9 | 74138 | 2 | AC021272 |
| 32 | 67 | 4.9 | 158167 | 2 | AC034105 |
| 33 | 66.8 | 4.9 | 62649 | 2 | AC022552 |
| 34 | 66.8 | 4.9 | 75628 | 2 | AC021793 |
| 35 | 66.8 | 4.8 | 100511 | 2 | AC010774 |
| 36 | 66.4 | 4.8 | 840 | 8 | CNS01BNS |
| 37 | 66.2 | 4.8 | 30168 | 2 | AC096156 |
| 38 | 66.2 | 4.8 | 61834 | 2 | AC068105 |
| 39 | 66.2 | 4.8 | 79355 | 2 | AC021280 |
| 40 | 66.2 | 4.8 | 82024 | 2 | AC023210 |
| 41 | 66.2 | 4.8 | 150695 | 2 | AC023576 |
| 42 | 66 | 4.8 | 66952 | 2 | AC105944 |
| 43 | 66 | 4.8 | 89994 | 2 | AC021263 |
| 44 | 66 | 4.8 | 134940 | 2 | AC018939 |
| 45 | 66 | 4.8 | 137878 | 2 | AP003300 |

ALIGNMENTS

RESULT 1
AX201324 LOCUS AX201324 1353 bp DNA
DEFINITION Sequence 3 from Patent WO0153486.
AX201324
VERSION AX201324
KEYWORDS GI:15391154
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1353)
AUTHORS Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L., Hillan,K.J., Marsters,S.A., Pan,J., Pletl,R.M., Roy,W.A., Smith,V., Stone,D.M., Watanabe,C.K. and Wood,W.I.
TITLE Compositions and methods for the treatment of tumour
JOURNAL Patent: WO 0153486-A 3 26-JUL-2001; Genentech, Inc. (US)
FEATURES
source location/Qualifiers
1..1353 /organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 257 a 443 c 389 g 264 t

ORIGIN

Query Match 96.2% Score 1320.2 DB 6 Length 1353;
Best Local Similarity 99.8% Pred. No. 2.8e-220;
Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Oy 49 cgaatccctcggatcccgagatgaggagggcggtgtagcagcagcagcccccccccacatg 108
Db 1 CGATCCCTCGGGTCCCGGGATGGGGGGCGGTAGAGCAGCAGCCACCCCGCCCGCAGG 60
Oy 109 gccgcgcgtcgagagccagagcgagagggcgcgcggggggagcgagcgagcgccctgctg 168
Db 61 GCCGCCCGCGTCGGAGCGCAGAGCGGAGGGGGCGCCGGGGGAGCCCGCGCCCTGTGCTG 120
Oy 169 gccgcgcgcgcgtgagcgctggcgctgagcgcgctgagcgctgagcgctgagcgctgag 228
Db 121 GTCCCGCTCGCGCTGGCGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGG 180
Oy 229 gtcagcttgaggagcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 288
Db 181 GTCACTTTGGGAGCGCGGGCATCGCTGTCCGCCAGAGACCTGCCCCAGAGAGAGCTGTG 240
Oy 289 gcaagagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 348
Db 241 GCAGAGAGAGAGCAGAGACCCCTCGGAACTGATCCCAAGACAGAAAGCCAGGATCCT 300
Oy 349 gccgccttcctcgaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 408
Db 301 GCGCCTTTCTTAAACCGACTAGTGTGGCTCGCAGAGAGTGCACCTTAAAGCCGGAAGAA 360
Oy 409 cggagctcgagagcgatcgagcagccatctgaagtcagcagcagcagcagcagcagcagcag 468
Db 361 CGGGCTCGAAGAGCATCGACGCCCATTTATGAAGTTCAATCCAGACCTGAGAGAGCGGA 420
Oy 469 gccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 528
Db 421 GCGCAGGCGAGGTGTGAGCGGAGCAGTGAATGCTGGGAGAGCAGAAATCAACAGCTCC 480
Oy 529 aagcctcgcgcgtacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 588
Db 481 AGCCCTCTGCGCTAACACCGCCAGATCGGGAGGATTATGATACCCGGGCTGGCTCTAC 540
Oy 589 taacctctcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 648
Db 541 TACCTGTACTGTACAGGTGACATTTGATGAGGGAAGGCTGTCTACTGTAAGCTGAGCTTG 600
Oy 649 ctggcggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 708
Db 601 CTGGTGGATGATGTGCTGTGCTGCGCTGCTGAGGAGAAATTCACAGCCACTGCGGGAGT 660
Oy 709 tcccccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 768
Db 661 TCCCTCGGGGCCCCAGCTCGCGCTCTGTCAGGTGTCTGGGCTGTGGCCCTGCGGCCAGGG 720
Oy 769 tccctccctcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 828
Db 721 TCCCTCGCTCGGATCGGACCGCTCCCTGCGGCCATCTCAAGGCTGCCCCCTTCTGACCC 780
Oy 829 taacctcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 888
Db 781 TACTTTCGAGCTTTTCCAGGTTACGTAGAGGGGCGCTGTGCTCCCGAGTGTCCAGAGCT 840
Oy 889 gccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 948
Db 841 GCGCGGCTCCCTCGAGAGCTCTCTGGGGCACCGGGTCCCTCTGCCCCACCTCTAGCGCGCT 900
Oy 949 ctcttgcacagacgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1008
Db 901 CTTTGTCTCAGACGTGCCCCCTCCCTCTAGAGGCTGCGCTGGGCTGTTCAGCGTGTTCACA 960
Oy 1009 tcccaataaataagatattcccaactctatattacaactcccccacgcgcacactcca 1068
Db 1009 tcccaataaataagatattcccaactctatattacaactcccccacgcgcacactcca 1068
```

```
Db 961 TCCACATTAATATACGATATTCCTACTTATCTTACAACTCCCCAGCCGACCTCTCCA 1020
Oy 1069 cctcaatagctccccaatccaccctgaccccttgaggcccccaatgactcgaactcccccctg 1128
Db 1021 CCTCACTAGCTCCCAATCCCTGACCCCTTTGAGGCCCCAGAGATCTGACCTCCCCCTG 1080
Oy 1129 gccacagaccccccaaggagcagctgctcactcagctcagctcagctcagctcagctcag 1188
Db 1081 GCCACAGACCCCGACAGCTATTGTGTCACTGTACTGTGAGGCAAGATGGGTCAGAGAAG 1140
Oy 1189 accccaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1248
Db 1141 ACCCCACTTTCAGGACCTAAGAGGGGCTGAGCTGCGCGCAGGACCAAGAGACTGGGC 1200
Oy 1249 ctaggcagagagctccccaatctgaggggcgaggaacaagaagcctccctcctgagaa 1308
Db 1201 CTAGGCCAGAGATTTCCCAATGTGAGGGCGAGAAACAGACACACTCTCCTTGAANA 1260
Oy 1309 ttcctctgagatctttaaacaagatatatttatttatttatttatttatttatttatttatt 1368
Db 1261 TTCCTGTGATTTTAAACAGATATTATTTTATTTATTTATTTATTTATTTATTTATTT 1320
Oy 1369 aatgg 1373
Db 1321 AATGG 1325
```

RESULT 2

AF055872 1368 bp mRNA linear PRI 04-MAY-1998

LOCUS AF055872 Homo sapiens Ap03/DR3 ligand (Ap03L) mRNA, complete cds.

DEFINITION AF055872
ACCESSION AF055872
VERSION AF055872.1 GI:3108230
KEYWORDS

ORGANISM human.

REFERENCE 1 (bases 1 to 1368)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS Ashkenazi, A., Sheridan, J.P., Pittl, R.M., Brush, J., Goddard, A. and

IDENTIFICATION OF A LIGAND FOR THE DEATH-DOMAIN-CONTAINING RECEPTOR

Ap03
Curr. Biol. 8 (9), 525-528 (1998)

JOURNAL MEDLINE 9822835
REFERENCE 2 (bases 1 to 1368)

AUTHORS Marshers, S.A., Sheridan, J.P., Pittl, R.M., Brush, J., Goddard, A. and

IDENTIFICATION OF A LIGAND FOR THE DEATH-DOMAIN-CONTAINING RECEPTOR

Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way,

South San Francisco, CA 94080, USA

Location/Qualifiers
1. 1368

FEATURES
source
1. 1368

1. 1368
/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="17p13"

1. 1368
/gene="Ap03L"

58. 807
/gene="Ap03L"

/function="binds to the death-domain-containing receptor

Ap03/DR3"

/note="NFE homolog"

/product="Ap03/DR3 ligand"

/protein_id="AAC39724.1"

/db_xref="GI:3108231"

/translational="MAARSSORRRGRRGPEPTALVPLATGLATLGLALLAVSL
GSRASLSAODEPAOEELVAEEDODPSLEPOTEEODPAFLNRLVRRPRAAPKGRKTR
ARRALIAHYEVHPRPDGDAQAGVGTYSWGEARINSPLRYNRIOIGETIVRAGI
YLVQVHFDEGSKAVYELKLDLLVQGVLAIRLLEERSATASLSLPQLRICOVSGLLAL
RPGSSLRIRLTPMAHLKAPFLTYFGLEOVH"

Db 361 GGACCGCAGGCGGTGTGACGGGACAGTGTGCTGGGAGGAACCAACAGC 420
Qy 526 tccagccctctgagctacaaccgagatcgaggatattatagacacggcgctg 585
Db 421 TCCAGCCCTTGCCCTACACACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTC 480
Qy 586 tactaccctgactgacagtgacacttgatagagggaaggtctgctactaaagctgac 645
Db 481 TACTACACTGTACTGTACAGTGTACTTTGATGAGGGAAGGCTGTACTAAGCTGAC 540
Qy 646 ttgctgtgtgagtgatgtgtgctgagccctgcctgctgaggaattctcagccatgagcc 705
Db 541 TTGCTGTGTGATGTGTGTCTGGCCCTGCGCTGAGGGAATTCAGCCACATGCGCGC 600
Qy 706 agttccctcgaggcccaagctccgctctgcaagtgatctgaggtgttgagccctcgagca 765
Db 601 AGTTCCCTCGGGCCCGCCAGCTCCGCTCTGCGAGGTGTGGGCTGTGGCCCTCGGCA 660
Qy 766 ggtctccctctgagctcgacacccctccctgagcccaatcacaagctgcccctctc 825
Db 661 GGGTCCCTCCCTGGGATCCGACCTCCCTCCGCGCATCTCAAGGCTGCCCTTCTC 720
Qy 826 acctacttggaactctctccaggtcactgagggccctgtgtctcccaagtgctccag 885
Db 721 ACCTACTTGGACCTTTCAGAGTTCAGTGAAGGGCCCTGCTCCCGCATGCTCCAG 780
Qy 886 gctgcccgtccctctgacaagctctctgggacacggctccctctgcccacccctgagcc 945
Db 781 GCTGCCCCGCTCCCTCAACACCTCTGTGGGACCCGGCTCCCTGCCCCACCTCAACC 840
Qy 946 gctcttgctcagaacctgcccctccctctagaagctgagccctgtgcttcaagtgctt 1005
Db 841 GCTCTTTGCTCCAACTGCCCCCTCCTCTAGAGGCTGCTGGGCTGTTACGTGTTT 900
Qy 1006 ccatccacataaatacagatttccactcttacttacaacatcccccacccactct 1065
Db 901 CCATCCCAACATTAATACAGTATTCCTCACTTATCTTACAACAACACCCACCCACTCT 960
Qy 1066 ccactactacagctcccaatccctgagcccttgaagcccccagtgatctgactcccc 1125
Db 961 CCACCTACTAGCTCCCAATCCCTGAACCTTTGAGGCCCCCACTGATCTGACTCCCC 1020
Qy 1126 ctggccacagaaccgagggatgtgtctcactgtactctgtggcaagatgagtcag 1185
Db 1021 CTGGCCACAGACCCCGAGGAGATGTGTCTACTCTGTGCGCAAGATGGGTCCAG 1080
Qy 1186 aagaccactcagcactaagaagggtgagccctgagccctgaggaagccaagaagactg 1245
Db 1081 AAGACCCCACTTCAAGCACTTAAGAGGGGCTGACCTGCGCGCAAGGCCAAAGAGACTG 1140
Qy 1246 ggcctagggcagaggttcccaatgtgaaggaggagaaacaagaacagctctcccttga 1305
Db 1141 GGCTTAGGCCAGAGGATTCCTCAATGTAGAGGGGAGAAAGCAAGACTCTCCCTTGA 1200
Qy 1306 gaattccctgtgatttttaaaacagatatattt 1341
Db 1201 GAATTCCCTGTGATTTTAAACAGATATTATTTT 1236

RESULT 5
BC019047 1651 bp mRNA linear PRI 11-DEC-2001
LOCUS BC019047
DEFINITION Homo sapiens, similar to tumor necrosis factor (ligand)
superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
complete cds.
ACCESSION BC019047
VERSION BC019047.1 GI:17512138
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1651)

AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: Louis Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgti.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Mastello, C., Mastian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stanthrop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgoun, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 30 Row: P Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA g1: 4507596.
FEATURES
source
1..1651
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:20669 IMAGE:4766071"
/tissue_type="Primary B-cells from Tonsils"
/clone_id="N1H-MGC_48"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
106..510
/codon_start=1
/product="Similar to tumor necrosis factor (ligand)
superfamily, member 12"
/protein_id="AAH19047.1"
/db_xref="GI:17512139"
/translation="MAARRSORRRGRGEGTLLVPLAIGLIALAGLLAVSL
GSRASUSAOEPADQELVAEDDPSELNPDYESODPAPFLNLRVRRSAPKGRKTR
ARRAIAHYEVHRRPGDGAADAGGYTCLRP"
CDS
BASE COUNT 344 a 517 c 481 g 309 t
ORIGIN
Query Match 79.9%; Score 1096.8; DB 9; Length 1651;
Best Local Similarity 84.8%; Pred. No. 2.2e-181;
Matches 1347; Conservative 0; Mismatches 2; Indels 239; Gaps 1;
Qy 25 tccgcccgcggctcccccctcccgatccctggtctcgagatgggggaggtgag 84
Db 25 TCCGCCCGCGGCTCCCTCCCTCCCGATCCCTCGGATGGGGGCGGTAGG 84
Qy 85 caggcacagccccccgcccccatagcgccgctggaagccagagcgagggcgccg 144
Db 85 CAGGCACAGCCCCCGCCCGCATGCGCCCGCTGGAGCCAGAGGCGAGGGCGCCGG 144
Qy 145 gggagccgggacacgcctctgtgtcccgctgagccctgagccctgagccctgagcc 204
Db 145 GGGAGCGCGGACACGCGCTGCTGTCGCCGCTCGGCGGCGCTGCGCGCTGAGCC 204
Qy 205 tgctctgagctctctgctgagcggtgtgtaattgggagacgggagatgcttccgcccag 264
Db 205 TGCTTCGGCTCTCTGCTGGCGGCTGTGAGTTGGGAGACCGGGATGCTTCCGCCAG 264
Qy 265 gagcctgcccagggagctgtgtgacagagagacccagccctggaactgactcc 324

```

Db 265 GAGCCCTGCCAGAGAGAGCTGGTGGAGAGGAGGACCCAGACCCGCTGGAACTGTAATCCC 324
Oy 325 cagacagaagaagaagccaggaatcctgcgccttcctgaacgcgaatctgagcttcgcaga 384
Db 325 cagacagaagaagaagccaggaatcctgcgccttcctgaacgcgaatctgagcttcgcaga 384
Oy 385 agtgcacctaagaagccaggaacacagcggctcgaagagcgaatcgcaacccatatagaatc 444
Db 385 agtgcacctaagaagccaggaacacagcggctcgaagagcgaatcgcaacccatatagaatc 444
Oy 445 celtccagacccctggagaagagcggagcagcagcagcagcagcagcagcagcagcagcagc 475
Db 445 cagtccacgacgctggacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 504
Oy 476 ----- 475
Db 505 CCATGAGATACTAGTGGTGGAGCCAGATTGAAACCCAGCTAGATGTGCTGTACT 564
Oy 476 ----- 475
Db 565 GGACATGCTCTCATGAAAGCGAGGTGGTGGCAGGGGTGAGGGGTCTCATGACAGGG 624
Oy 476 ----- 475
Db 625 CCACATCCAAAAAGGGAGAGGGAATTTCAGAAAGAGAGAGACACATCTCCACCA 684
Oy 476 ----- 475
Db 685 TTACCCAGAGGGTCAAAAGGAGAACAGAGACAGGTGTGACACGGAGAGTGAATGGCTGGG 744
Oy 506 aggaagccagaaatcaacagctccagcctctgcgtacacacccagcagatcggggagttta 565
Db 745 AGGAAGCCAGAAATCAACACACTCCAGCCCTGTGGGTACAAACCGCCAGATCGGGAGTTTA 804
Oy 566 tagtcaacccgggctgggctctactactctgtctgtcaggtgcacatttgatggggagag 625
Db 805 TAGTCAACCCGGGCTGGGCTCTACTACTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 864
Oy 626 cgtctacactgaagcctgcgtctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 685
Db 865 CTGTCTACTGAAAGCTGGAGCTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 924
Oy 686 aattctcaagcaactgcgcgcagcttcctctgcggccagcctcgcgcctctgcgcagtgctc 745
Db 925 AATTCTCAAGCACTGGGCGAGTTCCTCTGGGCGCCAGCTCCGCTCTGTCTGTCTGTCTGTCTGT 984
Oy 746 ggtctgtgctcctgcgcgcagaggtctcctcctgcgcgcagcctcctcctcctcctcctc 805
Db 985 GGTCTGTGGCTGGCGCGGCGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1044
Oy 806 tcaagagctcccccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 865
Db 1045 TCAAGGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1104
Oy 866 tctcccccagctcgtcccaagctgcgcgcctcctcctcctcctcctcctcctcctcctcctc 925
Db 1105 TCTCCCGGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1164
Oy 926 cctctgcgcgcacccctcagcgcctctgtctcagagcctcctcctcctcctcctcctcctcctc 985
Db 1165 CCTCTGCCACCCCTAGCGCGCTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1224
Oy 986 tgggcctgtcagctgttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1045
Db 1225 TGGGCTGTTCACGTCTTTTCCATCCACATAATACAGTATTTCCACCTTATCTTAA 1284
Oy 1046 actcccccacgcgcacactcctcactcactagctccccaatccctcctcctcctcctcctc 1105
Db 1285 ACTCCCGGACCGCCACTCTCCACCTCACTAGCTCCCAATCCCTGACCTTTGAGGCC 1344
Oy 1106 ccaagagatctgactcctccctcgcgcacagaccccccagggcattgtgtcactcgtactct 1165

```

```

Db 1345 CCAGTGAATCTGACTCCCTCCCTGGCCACAGACCCCGAGGGCAATGTCTACTGTACTCT 1404
Oy 1166 gtggcagaagatgggttccagaagaagccacactcagcgaactaagagggctgagctggcg 1225
Db 1405 GTGGCCAGAGATGGGTCCAGAGACCCCACTTCAGGCACTAAGAGGGGTGACCTGGCG 1464
Oy 1226 gcaagaagccaaagagactggcctagcagcagagttcccaaatgtgagggcgagaac 1285
Db 1465 GCAGAGAGCCAAAGAGAGACTGGCGCTAGGCGACAGAGATTCCTCAATGTGAGGGCGAGAAAC 1524
Oy 1286 aagaagactcctcctctgaagatccctcgtgagatctttaaagagatatatttatt 1345
Db 1525 AAGACAAAGCTCTCTCTTGAAGATTCTCTGTGATTTTAAACAGATATTTATTTATTT 1584
Oy 1346 attatgtacaaatgttgaataatgg 1373
Db 1585 ATTATTTGACAAATGTTGATTAATGG 1612

```

```

RESULT 6
AC016876 190358 bp DNA linear HTG 03-SEP-2001
LOCUS AC016876/c
DEFINITION Homo sapiens clone RP11-186B7, WORKING DRAFT SEQUENCE, 16 unordered
pieces.
ACCESSION AC016876
VERSION AC016876.4 GI:15421989
KEYWORDS HTG; PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 190358)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
TITLE Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL 2 (bases 1 to 190358)
UNPUBLISHED Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome, clone RP11-186B7

```

```

REFERENCE 1 (bases 1 to 190358)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
TITLE Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL 2 (bases 1 to 190358)
UNPUBLISHED Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome, clone RP11-186B7

```

```

TITLE
JOURNAL
COMMENT
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 13849
Center clone name: 186.B.7
Summary Statistics
Sequencing vector: M13; W77815; 3% of reads
Sequencing vector: Plasmid; n/a; 97% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 181928 bases at least Q40
Consensus quality: 185529 bases at least Q20
Consensus quality: 187335 bases at least Q20

```


RESULT 7
AX180714 898 bp DNA Linear PAT 06-AUG-2001
LOCUS Sequence 1 from Patent WO0145730.
DEFINITION AX180714
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 898
/organism="synthetic construct"
/db_xref="taxon:32630"
52. .873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=11
/protein_id="CAC5004.1"
/db_xref="GI:15132571"
/translation="MATGSRISLLAFGLCLPWLQEGSATSDRMKQIEDKIEILS
KIYHLENIAIRIKLIGERTSRSLASAEPAQELVAEEDQDSSELPTEES
ODPAFLNRLVRRPSAKGRKTRARIAAHVEHNPBGDQAGVYSGWEA
RINSSPLRYRQIGERTYRAGLYLYLCOVHNPDGKAVYLKDLNLYALRLIE
FSATRASSLGQDLRLCOVSGLLALPSSLSLRILPMAHLAAAPLTFGLFOVH"
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN

Query Match 45.8%; Score 629.2; DB 6; Length 898;
Best Local Similarity 99.5%; Pred. No. 5.9e-100;
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 232 agttgagagcagcgacatcgctgtccgcccagagccctgcagagagcgtgtgca 291
|||||
DB 250 attttggagagcggccgcatcgtctgtccgccagagacctgccagagagcgtgtgca 309
|||||
QY 292 gagagagcagagaccgtctcgaaactgaatcccccagagaaagccagatcctgcg 351
|||||
DB 310 gagagagcagagaccgtctcgaaactgaatcccccagagaaagccagatcctgcg 369
|||||
QY 352 ccttcctcgaacgactatctgagctcgacagaaagcaccctaaagagccgaaacacag 411
|||||
370 cctttcctgaacgactatctgagctcgacagaaagcaccctaaagagccgaaacacag 429
|||||

QY 412 gtcgaagagcgatcgacagcccatatgaagtcatccacagcctgcagagagcgagcg 471
|||||
DB 430 gctcgagagcgatcgacagcccatatgaagtcatccacagcctgcagagagcgagcg 489
|||||

QY 472 caggaagctgtggaagcagatgaatgagctgtggaagaaagccagaaatcaacagctccagc 531
|||||
DB 490 caggaagctgtggaagcagatgaatgagctgtggaagaaagccagaaatcaacagctccagc 549
|||||

QY 532 cctctcgagctaaacagcagatcgagaggttatagtcacccgggctgtgctctactac 591
|||||
DB 550 cctctcgagctaaacagcagatcgagaggttatagtcacccgggctgtgctctactac 609
|||||

QY 592 cgtactcgcagctgcaacttgaatgaagaggaagcgtctacactgaagctgagctgcg 651
|||||
DB 610 cgtactcgcagctgcaacttgaatgaagaggaagcgtctacactgaagctgagctgcg 669
|||||

QY 652 gtggaatgtgtgctgagcctgcgctgtggaagaaatctcagcaactgcggcagcttc 711
|||||
DB 670 gtggaatgtgtgctgagcctgcgctgtggaagaaatctcagcaactgcggcagcttc 729
|||||

QY 712 ctcgagagcccaagctcgagcctgtgcaaggtgtctggagctgtggcctgtgagcagagctcc 771
|||||
DB 730 ctcgagagcccaagctcgagcctgtgcaaggtgtctggagctgtggcctgtgagcagagctcc 789
|||||

QY 772 tccctgcgagctcgacacccctccctgggcccacatcgaagctgcacccctctcactac 831
|||||
DB 790 tccctgcgagctcgacacccctccctgggcccacatcgaagctgcacccctctcactac 849
|||||

QY 832 ttcgagactctccaggttcaactgaagggccctgcg 865
|||||
DB 850 ttcgagactctccaggttcaactgaagggccctgcg 883
|||||

RESULT 8
AF030100 1168 bp mRNA Linear ROD 20-DEC-1997
LOCUS
DEFINITION Mus musculus TWEAK mRNA, partial cds.
ACCESSION AF030100
VERSION AF030100.1 GI:2707220
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 1168
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="peritoneal macrophages"
<1. .679
/note="ligand in the TNF family; secreted protein"
/codon_start=2
/product="TWEAK"
/protein_id="AAC53517.1"
/db_xref="GI:2707221"
/translation="VLSGLALACILGLLVVSLGSMATLSAOPSOEILTAEDRREP
PELNQTEESODVPELFOILVRRPSAKGRARRAIAAHVEHNPBGDQAGV
DGVSGMEETKINSPLRYRQIGERTYRAGLYLYLCOVHNPDGKAVYLKDLN
GVLAIRCLIEFSATRASSPQDLRLCOVSGLLALPSSLSLRILPMAHLAAAPLTF
GLFOVH"
BASE COUNT 242 a 360 c 298 g 268 t
ORIGIN

Query Match 44.8%; Score 614.6; DB 10; Length 1168;
Best Local Similarity 75.8%; Pred. No. 2e-97;
Matches 906; Conservative 0; Mismatches 219; Indels 70; Gaps 9;

QY 178 ggcgtgagcctgagcctgtgagcctgtgagcctgtgagcctgtgagcctgtgagcctgtg 237
|||||
DB 2 gtcctgagcctgagcctgtgagcctgtgagcctgtgagcctgtgagcctgtgagcctgtg 61
|||||

QY 238 gggagccgggacatcgctgtccgcccagagacctgcagagagagctgtgagagagag 297
|||||
DB 62 gggagccgggacatcgctgtccgcccagagacctgtgagagagagctgtgagagagag 121
|||||

QY 298 gaccagagccgctcggaactgaatcccccagacagaagaagccagagatcctgcgcttc 357
|||||
DB 122 cgcctggagacccccctgaactgaatccccagacagaagaagccagagatcctgcgcttc 181
|||||

QY 358 ctgaaccgagctgagcctgcagaaatgacactaaagagccggagaaagcagagctcga 417
|||||
DB 182 ttggaaacaactgctcgcctcgagaaagctcctctaaagagccggagaaagcagagctcgc 241
|||||

[illegible]

| ACCESSION | LOCUS | DEFINITION | SEQUENCE | HTG | HTGS_PHASE1 | HTGS_DRAFT | HTGS_FULLTOP |
|--|------------|-------------------------------|---------------------------------|--------|-----------------|------------|--------------|
| AC069459 | AC069459/C | 203083 bp | DNA | linear | HTG_27-JUN-2001 | | |
| AC069459 | | Mus musculus chromosome 11 | clone RP23-168P5, WORKING DRAFT | | | | |
| AC069459 | | SEQUENCE, 7 unordered pieces. | | | | | |
| AC069459..23 | | GI:14547768 | | | | | |
| HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP | | | | | | | |

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 230883)
Metzker, M. L., Lewis, L. R., Hume, J., Edwards, C., Harris, C., Dederich, D., Thomas, S., Okunowo, G., Carlino, C., Garner, T., Addison, S., Pace, A., Williams, G., Bonini, D., Brooks, A., Brown, J., Butley, C., Bunick, C., Burkett, C., Chacko, J., Chen, G., Chen, Z., Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S., Fernandez, C., Ferraguto, D., Forcman-Tansey, J., Gill, R., Gorrell, J. H., Gunaratne, P., Haller, G., Hernandez, J., Hognes, M., Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S., Koval, C., Liu, J., Liu, W., Loulseged, H., Lozado, R. J., Martin, R., Massey, E., McLeod, M. P., Mei, G., Moore, S., Morgan, M., Morris, S., Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogih, M., Parish, B., Perez, L., Reller, D., Say, J., Shen, H., Vasquez, L., Watlington, S., Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A., Murthy, D. M., Rives, M., Scherer, S., Sodergren, E., Weinstock, G., Worley, K. and Gbbs, R.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 203083)
AUTHORS Worley, K. C.
TITLE Direct Submission
JOURNAL Submitted (31-MAY-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jun 25, 2001 this sequence version replaced gi:12621364.
----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: MAFO
Center clone name: RP23-168P5
----- Summary Statistics -----
Sequencing vector: M13; 108821
Chemistry: Dye-terminator Big Dye; 48% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 212648 bases at least Q40
Consensus quality: 218902 bases at least Q30
Consensus quality: 222384 bases at least Q20
Estimated insert size: 210656; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 7.2x in Q20 bases; sum-of-contigs estimation
----- NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html) -----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 62152: contig of 62152 bp in length
* 62153 62252: gap of unknown length
* 62253 118772: contig of 56520 bp in length
* 118773 118872: gap of unknown length
* 118873 148924: contig of 30052 bp in length
* 148925 149024: gap of unknown length
* 149025 167231: contig of 18207 bp in length
* 167232 167331: gap of unknown length
* 167332 189907: contig of 22576 bp in length
* 189908 190007: gap of unknown length
* 190008 196537: contig of 6530 bp in length
* 196538 196637: gap of unknown length
* 196638 203083: contig of 6446 bp in length.
Location/Qualifiers

source 1. 203083
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-168P5"
BASE COUNT 52662 a 49293 c 47892 g 52633 t 603 others
ORIGIN

Query Match 22.5%; Score 309.6; DB 2; Length 203083;
Best Local Similarity 72.0%; Pred. No. 1.7e-44;
Matches 556; Conservative 0; Mismatches 149; Indels 67; Gaps 9;

Db 601 caggtgcaattgtagaggaagcctgtctacccaagctgagctggtgagatggt 660
|||||
43019 CAGGTGCACTTGTAGAGGAAGAGGCTGCTACCTGAAGCTGACTGTGTGAAGCT 42960
|||||
661 gtgtcggccctggtgctgctggaagaattctcaagcactgagcaggtccctcgccc 720
|||||
42959 GTGTGCGCCCTGGCTGCTGGAGAAATTCAGCCACAGCAGCAAGCTCTCTGGGCC 42900
|||||
721 cagctcgccctctgcaaggtgtctggtgctgtgcccctgagggctccctctgag 780
|||||
42899 CAGCTCCGTTGTGCTCAGAGTGTCTGGGCTGTGGCCGCTGCGGCGAGGCTCTTCCG 42840
|||||
781 atccgacacccctccctgagccatctcaagctgcccctctctcaactacttcgagctc 840
|||||
42839 ATCCGACACCTCCCTGGGGCTATCTTAAGGCTGCCCCCTTCTTAACCTACTTTGAGACTC 42780
|||||
841 ttccaggtctcaatgtaggggcccctgtgtctcccaagtgctcccaagctgcgagctcc--- 897
|||||
42779 TTTCAAGTTTCATGAGGGGCTTGTCTCCAGATTCTTAACCTTTCCCTGGCTCAGAG 42720
|||||
898 ---cccgagcagctctctgagccagccggtccctctgcccacccctcaagcagctcttgtct 955
|||||
42719 AGCATCACCACACCTTCCTACCCCAACCCCACTCCTCCACCCCTTC- GCTGCTCCTGGT 42661
|||||
956 ccaagacctgcccctccctctcctagagctgctgagcctgtctcaagctgtttccatccaca 1015
|||||
42660 CCAGTCCGTCGTCCTCTCTC- AAAGGACAGCAGCAGCTTGTTACATGTTTTCATTCCACA 42602
|||||
1016 taaataagatctcccatctctatcttaataatcccccaagcccaactctcaactcaact 1075
|||||
42601 GA-----GCTATCTCTGCTCTTCTTAACATCCATCCACCAACAACATCCATCCACT 42547
|||||
1076 agctcccaatccctgagcccttgagccctcagtgatctcgaaccccccctgagcagag 1135
|||||
42546 AGCTCCCAAGCCCTACTATC-----CTGACCTCCCCACCACTCAC 42502
|||||
1136 acccccaagggactgtgtcactgtactctgtggaagagatgggtccagaagaccaccac 1195
|||||
42501 CCGACACAGCTGTTATTTGACTTTGTGCAC----- 42471
|||||
1196 ttcaagcacttaagagggctgtagcctgtagcaggaagcaagagactggcctagagcc 1255
|||||
42472 --CAGGCACTGAGATGGCTGGAACCTGGGAGCAAGAGCAAGCAAGCTGGGACTAGAGCC 42415
|||||
1256 aggaattcccaaatgtgagggcgagaa--aacaagacaagctcctcccttgaaatccct 1314
|||||
42414 AGAATTTCCCACTGTGAGGGGAGAGAGCTGGGACAAAGCTCTCTCCCTGGA----TCCCT 42359
|||||
1315 gtgagcttttaaaagagatatattatattatattatgtgacaaaatgttga 1366
|||||
42358 GTGGATTTTGAAA--AGATACATATTTTATTTATTTGACAAATATGTTAA 42309
|||||

RESULT 10
AL603707c 234182 bp DNA linear ROD 17-NOV-2001
LOCUS AL603707
DEFINITION Mouse DNA sequence from clone RP23-422L16 on chromosome 11,
complete sequence.
ACCESSION AL603707
VERSION AL603707.5 GI:17017790

KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 234182)
AUTHORS Pearce/A.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humanyes@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 20, 2001 this sequence version replaced gi:16605765.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; SW:
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-422L16 is
from the RP23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/dacpac/home.htm
VECTOR: pBAC3.6
This sequence is the entire insert of clone RP23-422L16.

FEATURES
LOCATION/Qualifiers
SOURCE 1. 234182
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_11b="RP23-422L16"
/clone_11b="RP23-422L16"
misc-feature complement(84050..84131)
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
BASE COUNT 59310 a 56824 c 57519 g 60529 t
ORIGIN

Query Match 22.5%; Score 309.6; DB 10; Length 234182;
Best Local Similarity 72.0%; Pred. No. 1.7e-44;
Matches 556; Conservative 0; Mismatches 149; Indels 67; Gaps 9;

Db 601 caggtgcaattgtagaggaagcctgtctacccaagctgagctggtgagatggt 660
|||||
66323 CAGGTGCACTTGTAGAGGAAGAGGCTGCTACCTGAAGCTGACTGTGTGAAGCT 66264
|||||
661 gtgtcggccctggtgctgctggaagaattctcaagcactgagcaggtccctcgagcc 720
|||||
66263 GTGTGCGCCCTGGCTGCTGGAGAAATTCAGCCACAGCAGCAAGCTCTCTGGGCC 66204
|||||
721 cagctcgccctctgcaaggtgtctggtgctgtgcccctgagggctccctcgagc 780
|||||
66203 CAGCTCCGTTGTGCTCAGAGTGTCTGGGCTGTGGCCGCTGCGGCGAGGCTCTTCCCTGCG 66144
|||||
781 atccgacacccctccctgagccatctcaaggtgtgcccctctctcaactacttcgagctc 840
|||||
66143 ATCCGACACCTCCCTGGGGCTATCTTAAGGCTGCCCCCTTCTTAACCTACTTTGAGACTC 66084
|||||
841 ttccaggtctcaatgtaggggcccctgtgtctcccaagctgtctcccaagctgcgagctcc--- 897
|||||
66083 TTTCAAGTTTCATGAGGGGCTTGTCTTCAGATTCTTCAACCTTTCCCTGGCTCAGAG 66024
|||||

| | | | |
|----|-------|---|-------|
| OY | 898 | --cctcgagcagctctcgtggaagccggtgcctctgcgcccacccctcaagcgcttctgt | 955 |
| Db | 66023 | AGCATACACCAACCTCCCTACCCACCCCACTCTCCACCCCTC-
GCTGCTCCTTGCT | 65965 |
| OY | 956 | ccagacgtccgcctcctcctctagaagctgacctgagcgctltcaagtttccatccaca | 1015 |
| Db | 65964 | CGAGTCCGTCTCTCC-
TTC-
AAGGCGACGACGAGCTTGTCAGTATGTTTCCATTCCACA | 65906 |
| OY | 1016 | taataacagtatctcccaactcttatcttacaacttcccccaacgcccactctccactctact | 1075 |
| Db | 65905 | GA-----GCTATCCTTGGCTCTCTTTAACTCCATCCACACCAACATATCCACTCACTCAT | 65851 |
| OY | 1076 | agctccccaatccccctgacaccttggagcccccagatcctgcagccccccctcgcgcaag | 1135 |
| Db | 65850 | AGCTCCCAAGCCCTACTTATC-----
CTGTACTTCCCCACCCACTTCAC | 65806 |
| OY | 1136 | accgccaggaactgtgtcttcaactgtaactctgtgagcaagatggtgccagaagaccacac | 1195 |
| Db | 65805 | CCGACACAGTGTTTATTTACTTTGTGCAC----- | 65775 |
| Y | 1196 | ctcaaggaactaagagggctggaacctgacgagcaagaacccaagagactgggacctagacc | 1255 |
| Db | 65776 | --CAAGGACCTAGAGAGGGCTGGACCTGGTGGCAGGAAGCCAGAGAACTGGGACTTAGGCC | 65719 |
| OY | 1256 | aggagttcccaaatgttggagggcgag-aacaagaacaagctctcccttgaagaattccct | 1314 |
| Db | 65718 | AGAAATTCACCAACTGTGAGGGGAGAGACTGGGGCAACAGCTCTCCTCTGGA---
TCCCT | 65663 |
| OY | 1315 | gtggatttttaaaacagatatattttttatatatttttggaaanaatttga | 1366 |
| Db | 65662 | GTGGATTGTGA--
AGATACTATTATTATTTATTTATGAGACAAATATGTAA | 65613 |

TITLE
JOURNAL
REFERENCES
AUTHORS
TITLE
JOURNAL
COMMENT

Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Mookentho, S.,
Oguri, M., Okwonu, G., Oragunye, N., Oriedro, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., L.,
Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojubokan, I., Rolfe, M.,
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshchari, N.,
Sisson, I., Sodegren, E., Sonalke, F., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, R., Teifrod, B., Thomas, J.,
Thomas, S., Umanh, K., Vazquez, L., Varra, V., Villalon, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,
Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 1790350)
Worley, K. C.

Direct Submission
Submitted (06-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17064677.

```
Assembly program: Phrap; version 0.990329First call to
findPhrapList
```

Consensus quality: 137096 bases at least Q40
Consensus quality: 145080 bases at least Q30
Consensus quality: 152950 bases at least Q20
Estimated insert size: 138980; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/centbank_draft_data.html)
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 68 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

| | | | |
|---|-------|-----------------------|--------------|
| * | 7467 | contig of 7467 | bp in length |
| * | 7468 | 7567: gap of unknown | length |
| * | 7568 | 12115: contig of 4548 | bp in length |
| * | 12116 | 12115: gap of unknown | length |
| * | 12216 | 20113: contig of 8098 | bp in length |
| * | 20314 | 20113: gap of unknown | length |
| * | 20414 | 25589: contig of 5176 | bp in length |
| * | 25590 | 25689: gap of unknown | length |
| * | 25690 | 31134: contig of 5745 | bp in length |
| * | 31435 | 31534: gap of unknown | length |
| * | 31535 | 37595: contig of 6061 | bp in length |
| * | 37596 | 37695: gap of unknown | length |
| * | 37696 | 42172: contig of 4477 | bp in length |
| * | 41173 | 42272: gap of unknown | length |
| * | 42273 | 46334: contig of 3962 | bp in length |
| * | 46335 | 46334: gap of unknown | length |
| * | 46335 | 50605: contig of 4271 | bp in length |
| * | 50606 | 50705: gap of unknown | length |
| * | 50706 | 55502: contig of 4497 | bp in length |
| * | 55203 | 55502: gap of unknown | length |
| * | 55303 | 58407: contig of 3105 | bp in length |
| * | 58408 | 58507: gap of unknown | length |
| * | 58508 | 61886: contig of 2779 | bp in length |

| | | |
|--------|---------|-----------------------------|
| 61387 | 61386: | gap of unknown length |
| 61387 | 61387: | contig of 3081 bp in length |
| 64468 | 64467: | gap of unknown length |
| 64468 | 64467: | gap of unknown length |
| 64498 | 64498: | contig of 3931 bp in length |
| 68499 | 68498: | gap of unknown length |
| 68599 | 68598: | gap of unknown length |
| 73450 | 73449: | contig of 4851 bp in length |
| 73550 | 73549: | gap of unknown length |
| 77402 | 77402: | contig of 3853 bp in length |
| 77403 | 77402: | gap of unknown length |
| 77503 | 81027: | contig of 3525 bp in length |
| 81128 | 81127: | gap of unknown length |
| 81128 | 81144: | contig of 3017 bp in length |
| 84245 | 84244: | gap of unknown length |
| 84245 | 86680: | contig of 2436 bp in length |
| 86681 | 86780: | gap of unknown length |
| 86781 | 89210: | contig of 2430 bp in length |
| 89211 | 89310: | gap of unknown length |
| 89311 | 929339: | contig of 3029 bp in length |
| 92340 | 92439: | gap of unknown length |
| 92440 | 94951: | contig of 2512 bp in length |
| 94952 | 95051: | gap of unknown length |
| 95052 | 95056: | contig of 3535 bp in length |
| 98567 | 98668: | gap of unknown length |
| 98667 | 100597: | contig of 1911 bp in length |
| 100588 | 100697: | gap of unknown length |
| 100698 | 102766: | contig of 2069 bp in length |
| 102767 | 102866: | gap of unknown length |
| 102867 | 105237: | contig of 2371 bp in length |
| 105238 | 105337: | gap of unknown length |
| 105338 | 107305: | contig of 1968 bp in length |
| 107306 | 107405: | gap of unknown length |
| 107406 | 109057: | contig of 1652 bp in length |
| 109058 | 109157: | gap of unknown length |
| 109158 | 111454: | contig of 2297 bp in length |
| 111455 | 111554: | gap of unknown length |
| 111555 | 113704: | contig of 2150 bp in length |
| 113705 | 113804: | gap of unknown length |
| 113805 | 115933: | contig of 2129 bp in length |
| 115934 | 116033: | gap of unknown length |
| 116034 | 118193: | contig of 2160 bp in length |
| 118194 | 118293: | gap of unknown length |
| 118294 | 120414: | contig of 2121 bp in length |
| 120415 | 120514: | gap of unknown length |
| 120515 | 122311: | contig of 1797 bp in length |
| 122312 | 122411: | gap of unknown length |
| 122412 | 126228: | contig of 3217 bp in length |
| 125629 | 125728: | gap of unknown length |
| 125729 | 128545: | contig of 2817 bp in length |
| 128546 | 128645: | gap of unknown length |
| 128646 | 130563: | contig of 1918 bp in length |
| 130564 | 130663: | gap of unknown length |
| 130664 | 133945: | contig of 2282 bp in length |
| 132966 | 133045: | gap of unknown length |
| 133046 | 134556: | contig of 1511 bp in length |
| 134557 | 134656: | gap of unknown length |
| 134657 | 136478: | contig of 1822 bp in length |
| 136479 | 136579: | gap of unknown length |
| 136579 | 138159: | contig of 1581 bp in length |
| 138160 | 138259: | gap of unknown length |
| 138260 | 140456: | contig of 2197 bp in length |
| 140457 | 140556: | gap of unknown length |
| 140557 | 142544: | contig of 1988 bp in length |
| 142545 | 142644: | gap of unknown length |
| 142645 | 143784: | contig of 1140 bp in length |
| 143785 | 143884: | gap of unknown length |
| 143885 | 145304: | contig of 1420 bp in length |
| 145305 | 145404: | gap of unknown length |
| 145405 | 147391: | contig of 1987 bp in length |
| 147392 | 147491: | gap of unknown length |
| 147493 | 148908: | contig of 1417 bp in length |
| 148909 | 149008: | gap of unknown length |
| 149009 | 150147: | contig of 1139 bp in length |
| 150148 | 150147: | gap of unknown length |

| Query Match | Best Local Similarity | Matches | Score | DB 2: | Length | DB 1: |
|---------------------------|---|-------------------------------------|--|--------|-----------------|-------|
| Query Match | Best Local Similarity | Matches | Score | DB 2: | Length | DB 1: |
| Matches 187; Conservative | 86.2%; Pred. No. 3,4e-18; | 0; Mismatches 29; Indels 1; Gaps 1; | | | | |
| 51 | atccctcgatcccggaaggggagcggtgagagacagcccccccgcccatgac | 110 | 11.6%; Score 158.6; DB 2: length 179030; | | | |
| 110928 | ATATGTTGGGTCCACAGATCATAGGCAGAGTTGCAGGCTCAGCCCCCGCCCATGG- | 110868 | | | | |
| 111 | cgcccgatcggaagcagagagcgaggggagcgccgagggagccgagccctgctggt | 170 | | | | |
| 110869 | CGCCGCTGCGAGACCAAGAGCGAGGAGGGGCGCGGGGAGCCGGGACCGCCCTGCTGAC | 110810 | | | | |
| 171 | cccgctcgagcgtcgagcgtcgagcgtcgagcgtcgagcgtcgagcgtcgagcgtggt | 230 | | | | |
| 110809 | CCCGCTGCTGCTGAGACCGCTGGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 110750 | | | | |
| 231 | cagttcgagagcagcgagcagtcgctcgcccgagag | 267 | | | | |
| 110749 | CAGCCTGGGAGACTGGGCACACGCTGTCTGCCACAGTG | 110713 | | | | |
| RESULT 12 | | | | | | |
| LOCUS | 166494 | 7218 bp | DNA | linear | PAT 28-DEC-1997 | |
| DEFINITION | Sequence 14 from patent US 5670367. | | | | | |
| ACCESSION | 166494 | | | | | |
| VERSION | 166494.1 | GI:2724471 | | | | |
| KEYWORDS | | | | | | |
| SOURCE | Unknown. | | | | | |
| ORGANISM | Unknown. | | | | | |
| REFERENCE | 1 (bases 1 to 7218) | | | | | |
| AUTHORS | Donner, F., Scheifflinger, F. and Falkner, F. Gunter. | | | | | |
| TITLE | Recombinant fowlpox virus | | | | | |
| JOURNAL | Patent: US 5670367-A 14 23-SEP-1997; | | | | | |
| FEATURES | Location/Qualifiers | | | | | |
| Source | 1..7218 | | | | | |
| BASE COUNT | 1944 a 1491 c 1486 g 1929 t 368 others | | | | | |
| ORIGIN | | | | | | |
| Query Match | 6.4%; Score 87.2; DB 6; Length 7218; | | | | | |
| Best Local Similarity | 5.5%; Pred. No. 1.2e-05; | | | | | |
| Matches 23; Conservative | 251; Mismatches 144; Indels 0; Gaps 0; | | | | | |
| 743 | ctggagctgtgagcctcgagcgagcgagctcctcctcgagatcccgacccctcgagcc | 802 | | | | |
| 1038 | CTTGGCTGAGGTCGAGGAGGAGCTTCGATATATATATATATATATATATATATATATAT | 1097 | | | | |
| 803 | atctcaagcgtgccccctccactactctgagactctcaggttcaagtgagggcc | 862 | | | | |
| 1098 | TT | 1157 | | | | |
| 863 | tggttccccaagctcgtccagcgtcgagcgtcctcctcagagctctctgggacccgg | 922 | | | | |

| | | | |
|----|------|--|------|
| Df | 1158 | yy | 1217 |
| Oy | 923 | tccctctgccccaccctcagcgctcttgcctccagaactgcccctccctataggct | 982 |
| Dd | 1218 | yy | 1277 |
| Oy | 963 | gctctggcctgttcacgtgtttccatccacaataaacaglatcccaactctattc | 1042 |
| Dd | 1278 | yy | 1337 |
| Oy | 1043 | aaactcccccccgcccaactctccaactcactagctcccaactccgaaccttgagg | 1102 |
| Dd | 1338 | yy | 1397 |
| Oy | 1103 | ccccccagatctcgaactccccctctgcacagacccccagggcatlgttgcactgt | 1160 |
| Dd | 1398 | yygAACCAATTCCTCATCTCT | 1455 |

RESULT 13

AF429315

LOCUS AF429315 125020 bp DNA linear PRI 18-JAN-2002

DEFINITION Homo sapiens junctophilin 3 (JPH3) gene, partial cds.

ACCESSION AF429315

VERSION AF429315.1 GI:17646244

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 125020)
Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)

TITLE

JOURNAL JOURNAL

MEDLINE 21583737

PUBMED 11694876

AUTHORS Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
Direct Submission
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA

JOURNAL TITLE Location/Qualifiers
1. 125020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16".
/map="16q24.3; between D16S520 and WI-12410"
/note="isolated from a patient with Huntington's
Disease-Like 2 (HDL2)"
complement(35581..35746)
/rpt_type= tandem
/rpt_unit= cty
complement(<36507..>36887)
/gene="JPH3"
/product="junctophilin 3"
complement(<36507..>36887)
/gene="JPH3"
/note="JPH3"
complement(<36507..36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/protein_id="AAI40941.1"
/db_xref="GI:17646245"
/translation="MSGSGRNPDDGSSYCGGWEDKRAHGVCITGRKGGEITYGMS
HGFEVLVTPSCNTTYOGTWAQKRKHIGLESYGKKWYTGEMTHGFRKGRGVRECAAG
NGAYVEGTWSMGIDGGYTELYSDG"

CDS

BASE COUNT 29056 a 32731 c 30696 g 28283 t 4254 others

ORIGIN

| | | | | |
|-----------------------|--------|--------------------|-----------------|----------------|
| Query Match | 5.7%; | Score 77.6; | DB 9; | Length 125020; |
| Best Local Similarity | 11.0%; | Pred. No. 0.00044; | | |
| Matches | 96; | Conservative 380; | Mismatches 396; | Indels 1; |
| | | | Gaps | 1. |

[illegible]

ACCESSION AC0084799
VERSION AC0084799.1 GI:11192127
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 303091)
REFERENCE DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 303091)
DOE Joint Genome Institute.
REFERENCE Submitted (17-NOV-2000) Production Sequencing Facility, DOE Joint
AUTHORS Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
JOURNAL
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 0
Center clone name: RPCT-23_197M9

Summary Statistics
Consensus quality: 152568 bases at least Q40
Consensus quality: 175579 bases at least Q30
Consensus quality: 186949 bases at least Q20
Estimated insert size: 198300; agarose-fp estimation
Estimated insert size: 293091; sum-of-contigs estimation
Quality coverage: 4.85 in Q20 bases; agarose-fp estimation
Quality coverage: 3.28 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 101 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1117 1116: contig of 1116 bp in length
1217 1216: gap of unknown length
2537 2536: contig of 1320 bp in length
2637 2636: gap of unknown length
4058 4057: contig of 1421 bp in length
4158 4157: gap of unknown length
5246 5245: contig of 1088 bp in length
5346 5345: gap of unknown length
6462 6461: contig of 1116 bp in length
6562 6561: gap of unknown length
7726 7725: contig of 1164 bp in length
7826 7825: gap of unknown length
9144 9143: contig of 1318 bp in length
9244 9243: gap of unknown length
10493 10492: contig of 1249 bp in length
10593 10592: gap of unknown length
11667 11666: contig of 1074 bp in length
11767 11766: gap of unknown length
12951 12950: contig of 1184 bp in length
13051 13050: gap of unknown length
14286 14285: contig of 1235 bp in length
14386 14385: gap of unknown length
15479 15478: contig of 1093 bp in length
15579 15578: gap of unknown length
16780 16779: contig of 1201 bp in length
16880 16879: gap of unknown length
18279 18278: contig of 1399 bp in length
18379 18378: gap of unknown length
19453 19452: contig of 1074 bp in length
19553 19552: gap of unknown length
21013 21012: contig of 1460 bp in length
21113 21112: gap of unknown length

21113 22169: contig of 1057 bp in length
22270 22269: gap of unknown length
23367 23366: contig of 1098 bp in length
23467 23466: gap of unknown length
24560 24559: contig of 1092 bp in length
24660 24659: gap of unknown length
25826 25825: contig of 1166 bp in length
25926 25925: gap of unknown length
27037 27036: contig of 1111 bp in length
27137 27136: gap of unknown length
28246 28245: contig of 1109 bp in length
28346 28345: gap of unknown length
29657 29656: contig of 1311 bp in length
29757 29756: gap of unknown length
30884 30883: contig of 1127 bp in length
30984 30983: gap of unknown length
32125 32124: contig of 1142 bp in length
32225 32224: gap of unknown length
33397 33396: contig of 1172 bp in length
33498 33497: gap of unknown length
34646 34645: contig of 1148 bp in length
34746 34745: gap of unknown length
35894 35893: contig of 1149 bp in length
35985 35984: gap of unknown length
35995 35994: gap of unknown length
37157 37156: contig of 1162 bp in length
37257 37256: gap of unknown length
38397 38396: contig of 1140 bp in length
38497 38496: gap of unknown length
39570 39569: contig of 1074 bp in length
39571 39570: gap of unknown length
40852 40851: contig of 1182 bp in length
40953 40952: gap of unknown length
42116 42115: contig of 1164 bp in length
42117 42116: gap of unknown length
43372 43371: contig of 1156 bp in length
43373 43372: gap of unknown length
43473 43472: gap of unknown length
44603 44602: contig of 1131 bp in length
44703 44702: gap of unknown length
44704 44703: contig of 1221 bp in length
45924 45923: gap of unknown length
46025 46024: gap of unknown length
47174 47173: contig of 1150 bp in length
47175 47174: gap of unknown length
47275 47274: gap of unknown length
48745 48744: contig of 1470 bp in length
48845 48844: gap of unknown length
49969 49968: contig of 1125 bp in length
49970 49969: gap of unknown length
50070 50069: gap of unknown length
51135 51134: contig of 1066 bp in length
51136 51135: gap of unknown length
51236 51235: gap of unknown length
52690 52689: contig of 1455 bp in length
52790 52789: gap of unknown length
52791 52790: gap of unknown length
53920 53919: contig of 1130 bp in length
54020 54019: gap of unknown length
55101 55100: contig of 1081 bp in length
55102 55101: gap of unknown length
55202 55201: gap of unknown length
56324 56323: contig of 1123 bp in length
56325 56324: gap of unknown length
57583 57582: contig of 1159 bp in length
57683 57682: gap of unknown length
58822 58821: contig of 1139 bp in length
58823 58822: gap of unknown length
58923 58922: gap of unknown length
60547 60546: contig of 1625 bp in length
60548 60547: gap of unknown length
60647 60646: gap of unknown length
61743 61742: contig of 1096 bp in length
61843 61842: gap of unknown length
62998 62997: contig of 1155 bp in length
63098 63097: gap of unknown length
64220 64219: gap of unknown length
64320 64319: contig of 1122 bp in length
64321 64320: gap of unknown length
65398 65397: contig of 1078 bp in length
65399 65398: gap of unknown length
66762 66761: contig of 1264 bp in length
66862 66861: gap of unknown length
80025 80024: contig of 1163 bp in length

| | | |
|---|-------|------------------------------------|
| * | 68026 | gap of unknown length |
| * | 68126 | 68125: gap of 1275 bp in length |
| * | 68101 | 69500: gap of unknown length |
| * | 69501 | 70718: contig of 1218 bp in length |
| * | 70719 | 70818: gap of unknown length |
| * | 70819 | 71995: contig of 1177 bp in length |
| * | 71996 | 72095: gap of unknown length |
| * | 72096 | 73232: contig of 1137 bp in length |
| * | 73233 | 73352: gap of unknown length |
| * | 73333 | 74476: contig of 1144 bp in length |
| * | 74477 | 74576: gap of unknown length |
| * | 74577 | 75863: contig of 1287 bp in length |
| * | 75864 | 75963: gap of unknown length |
| * | 75964 | 77021: contig of 1058 bp in length |
| * | 77022 | 77121: gap of unknown length |
| * | 77122 | 78212: contig of 1091 bp in length |
| * | 78213 | 78312: gap of unknown length |
| * | 78313 | 79424: contig of 1112 bp in length |
| * | 79425 | 79524: gap of unknown length |
| * | 79525 | 80617: contig of 1093 bp in length |
| * | 80618 | 80717: gap of unknown length |
| * | 80718 | 81804: contig of 1087 bp in length |
| * | 81805 | 81904: gap of unknown length |
| * | 81905 | 83007: contig of 1103 bp in length |
| * | 83008 | 83107: gap of unknown length |
| * | 83108 | 84187: contig of 1060 bp in length |
| * | 84168 | 84267: gap of unknown length |
| * | 84268 | 85347: contig of 1080 bp in length |
| * | 85348 | 85447: gap of unknown length |
| * | 85448 | 86521: contig of 1074 bp in length |
| * | 86522 | 86621: gap of unknown length |
| * | 86622 | 87744: contig of 1123 bp in length |
| * | 87745 | 87844: gap of unknown length |
| * | 87845 | 88965: contig of 1121 bp in length |
| * | 88966 | 89065: gap of unknown length |
| * | 89066 | 90176: contig of 1111 bp in length |
| * | 90177 | 90276: gap of unknown length |
| * | 90277 | 91279: contig of 1003 bp in length |
| * | 91280 | 91379: gap of unknown length |
| * | 91380 | 92563: contig of 1184 bp in length |
| * | 92564 | 92663: gap of unknown length |
| * | 92664 | 93688: contig of 1025 bp in length |
| * | 93689 | 93788: gap of unknown length |
| * | 93789 | 94934: contig of 1146 bp in length |
| * | 94935 | 95034: gap of unknown length |
| * | 95035 | 96111: contig of 1077 bp in length |
| * | 96112 | 96211: gap of unknown length |
| * | 96212 | 98352: contig of 2141 bp in length |
| * | 98353 | 98452: gap of unknown length |
| * | 98453 | 99645: contig of 1193 bp in length |
| * | 99646 | 99745: gap of unknown length |

[illegible][illegible]

| | |
|------------|---|
| RESULT | 15 |
| ACIO5667 | |
| LOCUS | ACIO5667 |
| DEFINITION | Rattus norvegicus clone CH230-12F12, *** SEQUENCING IN PROGRESS |
| ACCESSION | ACIO5667 |
| VERSION | ACIO5667.1 GI:18092890 |
| KEYWORDS | HTG; HTGS_PHASE1. |
| SOURCE | Norway rat. |
| ORGANISM | Rattus norvegicus |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; |
| | Rattus. |

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2002, 11:53:24 ; Search time 328.09 Seconds
(without alignments)
7184.990 Million cell updates/sec

Title: US-09-245-198a-3
Perfect score: 1373
Sequence: 1 atgcattgttgacttga.....gacaaatgtgataatg 1373

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_032802.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1 | 1373 | 100.0 | 1373 | 19 | AAV18600 |
| 2 | 1325 | 96.5 | 1421 | 20 | AAV56000 |
| 3 | 1320.2 | 96.2 | 1353 | 21 | AAV49717 |
| 4 | 1226.4 | 89.3 | 1236 | 19 | AAV47613 |
| 5 | 1226.4 | 89.3 | 1236 | 22 | AAV4350 |
| 6 | 833.4 | 60.7 | 1030 | 20 | AAV23424 |
| 7 | 629.2 | 45.8 | 898 | 22 | AAV30364 |
| 8 | 614.6 | 44.8 | 1168 | 19 | AAV18599 |
| 9 | 519.2 | 37.8 | 701 | 20 | AAV23425 |

| | | | | | | |
|------|-------|------|--------|----|----------|----------------------|
| 10 | 245.4 | 17.9 | 282 | 16 | AAV22190 | Human gene signal |
| C 11 | 70.4 | 5.1 | 114955 | 20 | AAV53491 | Human adenovirus A1 |
| 12 | 61 | 4.4 | 1337 | 20 | AAV17263 | Human gene express |
| C 13 | 60.6 | 4.4 | 1000 | 21 | AAV02484 | Human colon cancer |
| C 14 | 59.2 | 4.3 | 10732 | 21 | AAV10594 | Gene encoding a su |
| C 15 | 58 | 4.2 | 1218 | 21 | AAV02488 | Human colon cancer |
| C 16 | 54.4 | 4.0 | 1593 | 21 | AAV02504 | Human colon cancer |
| 17 | 53.2 | 3.9 | 53526 | 19 | AAV94101 | Human PKD1 gene. |
| 18 | 53.2 | 3.9 | 53577 | 17 | AAV18551 | Human polyomavirus k |
| 19 | 53.2 | 3.9 | 53577 | 19 | AAV94108 | Human PKD1 locus b |
| 20 | 50 | 3.6 | 50 | 20 | AAV56002 | Human tumour necro |
| 21 | 50 | 3.6 | 50 | 21 | AAV49732 | Human PRO207 DNA p |
| 22 | 50 | 3.6 | 201 | 8 | AAV70195 | Streptomyces prote |
| 23 | 50 | 3.6 | 2188 | 20 | AAV27506 | Human ovarian tumo |
| 24 | 49.6 | 3.6 | 1517 | 21 | AAV20208 | CDNA encoding huma |
| 25 | 49.2 | 3.6 | 825 | 20 | AAV25378 | Human mFLINT #2 nu |
| 26 | 49.2 | 3.6 | 825 | 21 | AAV51078 | Human mature FLINT |
| 27 | 49.2 | 3.6 | 936 | 20 | AAV25376 | Human FLINT #2 nuc |
| 28 | 49.2 | 3.6 | 936 | 21 | AAV88731 | Human FAS ligand 1 |
| 29 | 49.2 | 3.6 | 936 | 21 | AAV51076 | Human FLINT coding |
| 30 | 49.2 | 3.6 | 936 | 22 | AAV07385 | Human fas ligand 1 |
| 31 | 49.2 | 3.6 | 936 | 22 | AAV89920 | Nucleotide sequenc |
| 32 | 49.2 | 3.6 | 936 | 22 | AAV84738 | Nucleotide sequenc |
| 33 | 49.2 | 3.6 | 936 | 22 | AAV84739 | Nucleotide sequenc |
| 34 | 49.2 | 3.6 | 936 | 22 | AAV7696 | Human FLINT native |
| C 35 | 49 | 3.6 | 1126 | 21 | AAV02538 | Human colon cancer |
| C 36 | 48.8 | 3.6 | 201 | 8 | AAV70194 | Signal portion of |
| C 37 | 48 | 3.5 | 320 | 21 | AAV38185 | Primer used in the |
| 38 | 48 | 3.5 | 109519 | 22 | AAV50863 | Microsomopora DNA |
| 39 | 48 | 3.5 | 114955 | 20 | AAV53491 | Human adenovirus A1 |
| 40 | 47.6 | 3.5 | 767 | 20 | AAV23419 | Human APO6 DNA. H |
| 41 | 47.6 | 3.5 | 813 | 20 | AAV25377 | Human mFLINT #1 nu |
| 42 | 47.6 | 3.5 | 813 | 21 | AAV75989 | DNA encoding a mat |
| 43 | 47.6 | 3.5 | 813 | 21 | AAV88730 | Human FAS ligand 1 |
| 44 | 47.6 | 3.5 | 813 | 21 | AAV51077 | Mature human FLINT |
| 45 | 47.6 | 3.5 | 813 | 22 | AAV07380 | Human mature fas 1 |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AAV18600 | standard; CDNA; 1373 BP. |
| ID | AAV18600; |
| XX | AAV18600; |
| AC | 21-JUN-1998 (first entry) |
| XX | |
| DT | |
| XX | |
| DE | Human sapiens tumour necrosis factor related ligand (TNFRL) gene. |
| XX | |
| KW | TNFR1; tumour necrosis factor related ligand; tnfr; treatment; |
| KW | cancer; autoimmune disease; immune system; stimulation; suppression; |
| KW | graft rejection; ds. |
| XX | |
| OS | Human sapiens. |
| XX | |
| FT | |
| FT | key |
| FT | 1..852 |
| FT | Location/Qualifiers |
| PN | /note= "tumour necrosis factor related ligand" |
| XX | |
| PD | W09805783-A1. |
| PD | 12-FEB-1998. |
| XX | |
| PE | 07-AUG-1997; 97WO-US13945. |
| XX | |
| PR | 18-MAR-1997; 97US-0040820. |
| PR | 07-AUG-1996; 96US-0023541. |
| PR | 18-OCT-1996; 96US-0028515. |
| XX | |
| PA | (BIOL) BIOGEN INC. |

| | |
|----|---|
| OS | Homo sapiens. |
| XX | |
| FH | Key |
| FT | CDS |
| FT | |
| FT | |
| XX | |
| PN | W09919490-A1. |
| XX | |
| PD | 22-APR-1999. |
| XX | |
| PF | 09-OCT-1998; 98WO-US21407. |
| XX | |
| PR | 17-DEC-1997; 97US-0069862. |
| PR | 10-OCT-1997; 97US-0062037. |
| XX | |
| PA | (GETH) GENENTECH INC. |
| XX | |
| PI | Ashkenazi AJ, Marsters SA, Pittl R; |
| XX | |
| R | WPI: 1999-287982/24. |
| RR | P-PSDB; AAY09369. |
| XX | |
| PT | New human Apo-3 ligand (a tumour necrosis factor) homologue |
| XX | |
| PS | Clatm 18; Fig 1; 74pp; English. |
| XX | |
| CC | The present sequence encodes a human tumour necrosis factor (TNF) and |
| CC | lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has |
| CC | cyclostatic activity. Apo-3 ligand can be used to induce apoptosis in |
| CC | mammalian cancer cells, to induce NF-kappab-dependent transcription and |
| CC | to induce JNK/SAPK-dependent responses in mammalian cells. |
| XX | |
| SO | Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other; |

| | | | | |
|-----------------------|--------------|---------------------|----------------|--------------|
| Query Match | 96.5%; | Score 1325; | DB 20; | Length 1421; |
| Best Local Similarity | 98.5%; | Pred. No. 8.8e-273; | | |
| Matches 1337; | Conservative | 0; | Mismatches 20; | Indels 0; |
| | | | | Gaps 0; |

[illegible]

| | | | |
|---|--------------|---|-------|
| Db | 483 | gTgcgtcggaggaagccgaatacaacagctcccaagccctctgcgtctacaacccgcagatcg | 54.2 |
| OY | 557 | gggaattctatagtcacccgggctcgggcctactacacgttactgtcaggttcattatg | 61.8 |
| Db | 543 | gggaattctatagtcacccgggctcgggcctactacacgttactgtcaggttcattatg | 60.2 |
| OY | 617 | aggggaaagcgtctaccctgaagctbgaactctgtctgtgtgaTgtgtctgTggccctgcgt | 67.6 |
| Db | 603 | aggggaaagcgtctaccctgaagctbgaactctgtctgtgtgaTgtgtctgTggccctgcgt | 66.2 |
| OY | 677 | gccttgaaagaaattctcaagccactgcgcgaagtcccttcggggcccaagctccgcctgcg | 73.6 |
| Db | 663 | gccttgaaagaaattctcaagccactgcgcgaagtcccttcggggcccaagctccgcctgcg | 72.2 |
| OY | 737 | aggTgtctcggggctgtgtggccctgcgcgaaggtctccctctcggatctgcgaacccctccct | 73.6 |
| Db | 723 | aggTgtctcggggctgtgtggccctgcgcgaaggtctccctctcggatctgcgaacccctccct | 72.2 |
| OY | 797 | ggggccatctcaagcgtctccccccttcctcaactactaattcggactcttcacagttcaactgag | 85.6 |
| Db | 783 | ggggccatctcaagcgtctccccccttcctcaactactaattcggactcttcacagttcaactgag | 84.2 |
| OY | 857 | ggggccgtgtctccccaagtcgtcccaagctgcgcggtctccctctcgacaagctctctgggc | 91.8 |
| Db | 843 | ggggccgtgtctccccaagtcgtcccaagctgcgcggtctccctctcgacaagctctctgggc | 90.2 |
| OY | 917 | accgggtccctctctgcgcccaacctcgaagcgtctcttctgcagaacgtgcgcctctcccta | 97.6 |
| Db | 903 | accgggtccctctctgcgcccaacctcgaagcgtctcttctgcagaacgtgcgcctctcccta | 96.2 |
| OY | 977 | gaggtctgcctggcgctgtctctacagTgtttccatcccaataataatacagtatcccaactc | 103.3 |
| Db | 963 | gaggtctgcctggcgctgtctctacagTgtttccatcccaataataatacagtatcccaactc | 102.2 |
| OY | 1037 | tatcttacaactccccaacgcgcgcacactctccacactcaagctctcccaatccctgaccc | 109.9 |
| Db | 1023 | tatcttacaactccccaacgcgcgcacactctccacactcaagctctcccaatccctgaccc | 108.8 |
| OY | 1097 | ttgaaagcccccaagtgatctgtgaactcccccctgggcgaagaagcccccaaggtattgttca | 115.5 |
| Db | 1083 | ttgaaagcccccaagtgatctgtgaactcccccctgggcgaagaagcccccaaggtattgttca | 114.4 |
| OY | 1157 | ctgtactctctgtggcgaaagatctgtgtccagaagaccaccaactcagaagcctaagaagggctg | 121.1 |
| Db | 1143 | ctgtactctctgtggcgaaagatctgtgtccagaagaccaccaactcagaagcctaagaagggctg | 120.2 |
| OY | 1217 | gacctgtgcggcagaagcgaagaagactgggcctaaggccagaagattcccaatgtgtgaagg | 127.7 |
| Db | 1203 | gacctgtgcggcagaagcgaagaagactgggcctaaggccagaagattcccaatgtgtgaagg | 126.6 |
| OY | 1277 | gcgagaagaacaagacaaagctctccctcttgagaatactccctctgtgattctttaaacagatatc | 133.3 |
| Db | 1263 | gcgagaagaacaagacaaagctctccctcttgagaatactccctctgtgattctttaaacagatatc | 132.2 |
| OY | 1337 | attcttatattatgtgtacaacaattgtgtgataatbg | 137.3 |
| Db | 1323 | attcttatattatgtgtacaacaattgtgtgataatbg | 135.9 |
| RESULT 3 | | | |
| AAAA9717 | | | |
| ID | AAA49717 | standard; cDNA; 1353 BP. | |
| AAA49717; | | | |
| AC | XX | | |
| XX | XX | | |
| DT | 25-SEP-2000 | (first entry) | |
| XX | XX | | |
| DE | Human PRO207 | cDNA clone DNA30879-1152. | |
| PRO207; human; antitumour; tumour; therapy; cytostatic; breast cancer; ovarian cancer; renal cancer; colorectal cancer; uterine cancer; prostate cancer; lung cancer; bladder cancer; | | | |
| KM | | | |

| | | | |
|----------------------------|---|---------------|--|
| KW | | | central/nervous system cancer; melanoma; leukaemia; neoplasia; ss. |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| Key | Location/Qualifiers | | |
| FT CDS | /tag= a
58..807 | | |
| FT sig_peptide | /tag= b
58..177 | | |
| FT mat_peptide | /tag= c
178..804 | | |
| FT | | | |
| XX | | | |
| PV | WO200037638-A2. | | |
| PD | 29-JUN-2000. | | |
| XX | | | |
| PE | 02-DEC-1999; | 99MO-US28565. | |
| XX | | | |
| PR | 22-DEC-1998; | 98US-O113296. | |
| PR | 08-MAR-1999; | 99MO-USO5028. | |
| PR | 21-APR-1999; | 99US-O130232. | |
| PR | 28-APR-1999; | 99US-O131445. | |
| PR | 14-MAY-1999; | 99US-O134287. | |
| PR | 20-JUL-1999; | 99US-O144758. | |
| PR | 26-JUL-1999; | 99US-O145698. | |
| PR | 15-SEP-1999; | 99MO-US21090. | |
| PR | 15-SEP-1999; | 99MO-US21547. | |
| XX | | | |
| PA | (GETH) GENENTECH INC. | | |
| XX | | | |
| PI | Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA, | | |
| PI | Napier MA, Pittl RM, Wood WI; | | |
| XX | | | |
| DR | MP1: 2000-442668/38. | | |
| DR | P-PDB: AAU95338. | | |
| PT | Novel composition to inhibit neoplastic cell growth or for treating tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219, PRO221, PRO324, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866 - | | |
| PS | Claim 20; Fig 3; 172pp; English. | | |
| CC | The present sequence is that of CDNA clone DNA30879-1152 (ATCC 209358) encoding human PRO207 (see AAU95338), which shows homology to several members of the tumour necrosis factor family, especially human lymphotoxin (23.4%). The CDNA was identified in a foetal kidney cDNA library following identification of an expressed sequence tag with homology to human Apo-2 ligand. A claimed method for inhibiting the growth of a tumour cell comprises exposing the tumour cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO324, PRO328, PRO301, PRO526, PRO362, PRO509 or PRO866 (see AAU95337-49), their agonists or chimeric polypeptides incorporating them. The tumour is especially a cancer selected from breast, ovarian, renal, colorectal, uterine, prostate, lung, bladder and central nervous system cancer, melanoma and leukaemia. Nucleic acids encoding PRO179 etc. are used in the recombinant production of the antitumour polypeptides. | | |
| SQ | Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other: | | |
| Query Match | 96.2%; Score 1320.2; DB 21; Length 1353;
Best Local Similarity 99.8%; Pred. No. 9.2e-272; | | |
| Matches 1322; Conservative | 0; Mismatches 3; Indels 0; Gaps 0; | | |
| OY | 49 gcattccctcggttcgccgagatggagggcggcgggaacagcacaccgcccccccatg 108 | | |
| Dd | 1 cgattccctcggttcgccgagatggagggcggcgggaacagcacaccgcccccccatg 60 | | |
| OY | 109 gccgcgcgttccgagccagaagcggcgaaggcggcggcgggaacccgcgcctctgcy 168 | | |
| Dd | 61 gccgcgcgttccgagccagaagcggcgaaggcggcggcgggaacccgcgcctctgcy 120 | | |

[illegible]

```

Oy      1249 ctaggccagcaggttcccgaatgtagggcgagaaacagaacagtctcccttgagaa   1308
Db      1201 ttatttttgttatgttcaccataatgtgaggggcgaaaacaagaacagctccctcttgagaa   1260
Oy      1309 ttcctctgtgatTTTtaaacagatatTTTttattattatGTGACAaatgttgata   1368
Db      1261 ttcctctgtgatTTTTtaaacagatatTTTttattattatGTGACAaatgttgata   1320
Oy      1369 aaatg 1373
        |||||
Db      1321 aaatg 1325

RESULT . 4
AAV47613 ID AAV47613 standard; cDNA; 1236 BP.
XX AA V47613 :
XX AC
XX CT 27-Oct-1998 (first entry)
XX AT
XX GC
XX TT TNF related endothelium proliferative agent gene.
XX CC ss; TNF; endothelium proliferative agent; TREPA: wound healing; cancer;
XX KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FH CDS 1..750
FT FT /*tag= a
FT FT /product= "TREPA"
XX XX
XX PN MO9835061-AZ.
XX PD 13-Aug-1998.
XX PF 12-FEB-1998; 98WO-US02859.
XX PR 10-FEB-1998; 98US-0021706.
XX PA 12-FEB-1997; 97US-0798692.
XX PA (ABBO ) ABBOTT LAB.
XX PI WILEY SR:
XX DR MPI: 1998-AA7255/38.
XX DR P-PSDB; AAM29745.
XX XX
XX XX Detecting nucleic acid encoding TREPA - useful for diagnosis and
XX XX treatment of autoimmune disease, tumours and inflammation
XX XX
XX XX Claim 11: Page 123-4; 142pp; English.
XX XX
XX XX The TNF-related endothelium proliferative agent (TREPA), or its
XX XX activators or agonists, are used to treat a deficit of TREPA, e.g. to
XX XX promote wound healing or tissue grafting, by promoting vascularisation,
XX XX also to induce apoptosis for treating cancer and eliminating autoreactive
XX XX T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
XX XX TREPA peptides can also be used to target cytotoxic agents or for
XX XX affinity isolation of the corresponding receptor, the nucleic acid for
XX XX which can be used to transform tumour cells to render them more
XX XX responsive to TREPA and to screen for TREPA mimics.
XX XX Ribozymes, antisense RNA , antibodies or peptides, are used to treat
XX XX TREPA-associated diseases, e.g. tumours and metastases (by inhibiting
XX XX vascularisation), inflammation or a wide range of autoimmune conditions,
XX XX conditions involving abnormal stimulation of epithelial cells (e.g.
XX XX atherosclerosis), for birth control (inhibiting ovulation and placental
XX XX formation) or other angiogenic conditions (e.g. ulcers).
XX XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;
```

[illegible]

```
|||||
Db 1021 ctggccacagagaccgccagggcattgltcactgtactctgtggcaggaagggtccag 1080
Oy 1186 aagaccacacttcagcactaagagagcctgagcctggcgaggaagccaaagagactg 1245
Db 1081 aagaccacacttcagcactaagagagcctgagcctggcgaggaagccaaagagactg 1140
Oy 1246 ggcctagcagcagagattcccaaatgtgagggcgagaaacaagaagctcctcttga 1305
Db 1141 ggcctagcagcagagattcccaaatgtgagggcgagaaacaagaagctcctcttga 1200
Oy 1306 gaattccctgtgattttttaaacaagatatattt 1341
Db 1201 gaattccctgtgattttttaaacaagatatattt 1236

RESULT 5
AADD04350
ID AADD04350 standard: cDNA: 1236 BP.

AADD04350:
04-JUL-2001 (first entry)

DE Human TREPA (TNF related endothelium proliferative agent) cDNA.
KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vulnereary; ss.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..750
XX FT /*tag= a
XX FT /product= "Human TREPA (TNF related endothelium
XX FT proliferative agent)"
XX
XX US6207642-B1.
XX
XX 27-MAR-2001.
XX
XX 26-JUN-1998; 98US-0105343.
XX
XX 12-FEB-1997; 97US-0798692.
XX 10-FEB-1998; 98US-0021706.
XX
XX (ABBO ) ABBOTT LAB.

Wiley SR:
XX
DR WPI: 2001-280760/29.
DR P-PSDB: AAE00891.
XX
XX Inducing angiogenesis in mammal at desired sites for promoting wound
XX healing, by administering soluble fragment of extracellular domain of
XX tumor necrosis factor related endothelium proliferative agent protein
XX
XX Example 2: Column 73-74: 53pp: English.
XX
XX The present invention relates to extracellular signal molecules,
XX particularly members of tumour necrosis factor (TNF) family molecules
XX designated as TREPA (TNF related endothelium proliferative agent).
XX Soluble biologically active TREPA are used to treat TREPA-associated
XX diseases, tumours or metastases. TREPA is used for inducing angiogenesis
XX in human for promoting wound healing and for vascularising grafted
XX tissue for successful grafting and to promote tissue grafts.
XX The present sequence is a cDNA clone ID #690050 encoding human TREPA.
XX
XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other.
```

```
Query Match 89.3%, Score 1226.4; DB 22; Length 1236;
Best Local Similarity 99.5%; Pred. No. 8.4e-252;
Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 106 atggcgcgcgcgtcgagacagagagcgagggcgcgcgaggagccggagccggccctg 165
Db 1 atggcgcgcgcgcgtcgagacagagagcgagggcgcgcgaggagccggagccggccctg 60
Oy 166 ctgtgtccgcctcgtcgtctggccttgagccttgagccttgagccttcctcgtctgagcc 225
Db 61 ctgtgtccgcctcgtcgtctggccttgagccttgagccttgagccttcctcgtctgagcc 120
Oy 226 gtgtcagttcgtggagacgcggcagtcgctgtcgcgcagagacgttgcacagagagctg 285
Db 121 gtgtcagttcgtggagacgcggcagtcgctgtcgcgcagagacgttgcacagagagctg 180
Oy 286 gtgtcagagagagagacagagacccgtcggaaactgaaatcccccagacagaaagccagat 345
Db 181 gtgtcagagagagagacagagacccgtcggaaactgaaatcccccagacagaaagccagat 240
Oy 346 cctgcgccttcctcgtacacagactagttcgcgcctgcagagaatgacacccaagccgagaa 405
Db 241 cctgcgccttcctcgtacacagactagttcgcgcctgcagagaatgacacccaagccgagaa 300
Oy 406 acaacggctcgaagaagcgatcgcagcccatatgaaatcatalccacgaccccgacagagac 465
Db 301 acaacggctcgaagaagcgatcgcagcccatatgaaatcatalccacgaccccgacagagac 360
Oy 466 ggaagcagagcagaggtgtgtgacggagacagtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 525
Db 361 ggaagcagagcagaggtgtgtgacggagacagtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
Oy 526 tccagccctcgtcgtacacacccgcagatcggagaggttataagtaaccggcgtgtgtc 585
Db 421 tccagccctcgtcgtacacacccgcagatcggagaggttataagtaaccggcgtgtgtc 480
Oy 586 tactactgttactgtcaggtgtcacttgaatgaggggaagcgtgtactcgaagcgtgagc 645
Db 481 tactactgttactgtcaggtgtcacttgaatgaggggaagcgtgtactcgaagcgtgagc 540
Oy 646 ttgcctggtggaatgtgtcgtgcctcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 705
Db 541 ttgcctggtggaatgtgtcgtgcctcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 600
Oy 706 agttccctcggcccccaagctccgcctcgtgcagagtgatctgtgtgtgtgtgtgtgtgtgt 765
Db 601 agttccctcggcccccaagctccgcctcgtgcagagtgatctgtgtgtgtgtgtgtgtgtgt 660
Oy 766 gggctccctcgtcgtggtcgcacccctccctgtggcgcacatcgaagctgtccctcctc 825
Db 661 gggctccctcgtcgtggtcgcacccctccctgtggcgcacatcgaagctgtccctcctc 720
Oy 826 acctactcgaactcttcgaagttactgagggagccgtgctccacgaatctgtccag 885
Db 721 acctactcgaactcttcgaagttactgagggagccgtgctccacgaatctgtccag 780
Oy 886 gctgcgcgctccctcgcagacgtctctgtggcacccgcgtccctcgtccacacccctcagcc 945
Db 781 gctgcgcgctccctcgcagacgtctctgtggcacccgcgtccctcgtccacacccctcagcc 840
Oy 946 gctcttgcctcagacgtgcctccctcctcctagaggtcctgtggcctgttcaagtgttt 1005
Db 841 gctcttgcctcagacgtgcctccctcctcctcctcctagaggtcctgtggcctgttcaagtgttt 900
Oy 1006 ccatccacataaatacagatctccacatctatcttcaactcccccacggccacact 1065
Db 901 ccatccacataaatacagatctccacatctatcttcaactcccccacggccacact 960
Oy 1066 ccaactcctagctccccaatccctgaccccttggagcccccagtgatcgaactccccc 1125
Db 961 ccaactcctagctccccaatccctgaccccttggagcccccagtgatcgaactccccc 1020
Oy 1126 ctggccacagaccccccagggcattgttcaactgtactcgtgtggcagaagatgggtccag 1185
```

|||||
Db 1021 ctggcacaagaccgccagggatgctgtctactctgtggaaggttggtccag 1080
Qy 1186 aagccccacttcagcactaagaggggctgacctgagcgcaagaagccaaagagactg 1245
Db 1081 aagccccacttcagcactaagaggggctgacctgagcgcaagaagagagactg 1140
Qy 1246 ggcctagagccagaggttccccaatctgagggcgagaacaagaagcctctcccttga 1305
Db 1141 ggcctagagccagaggttccccaatctgagggcgagaacaagaagcctctcccttga 1200
Qy 1306 gaattccctgtggaattttaaacaagatatattt 1341
Db 1201 gaattccctgtggaattttaaacaagatatattt 1236

RESULT 6
AAK23424
ID AAK23424 standard; DNA: 1030 BP.
XX
XX AAK23424:
DT 18-JUN-1999 (first entry)
XX
XX Human TNRL3 DNA.
DE
XX Tumour necrosis factor receptor: signal transducer molecule: TNF; APO4;
KM developmental abnormality; gestational abnormality; prostate cancer;
KM APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
XX apoptosis; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..627
FT /*tag= a
FT /product= "TNRL3"
XX
XX MO9911791-A2.
XX
XX 11-MAR-1999.
PD
XX
PF 04-SEP-1998; 98MO-US181393.
XX
XX 05-SEP-1997; 97US-0924634.
PR
XX (UNIW) UNIV WASHINGTON.
PA
XX Chaudhary PM.
XX
XX WPI: 1999-205191/17.
DR P-PSDB; AAW93590.
XX
XX
XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX
XX Example VII; Fig 13A; 156bp; English.
XX
XX This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/ active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in APO4
CC activity. Effective pharmacological agents useful in diagnosis or

CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other:

Query Match 60.7%; Score 833.4; DB 20; Length 1030;
Best Local Similarity 99.9%; Pred. No. 3.7e-168;
Matches 834; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 229 gtcaatttgaggagccgggcatcgtctcgcccaagagcctgcgcaagagagctggtg 288
Db 1 gtcaatttgaggagccgggcatcgtctcgcccaagagcctgcgcaagagagctggtg 60
Qy 289 gcagagagagagccagagccgctcggaactgaatccccaagagaagaagccagatcct 348
Db 61 gcagagagagagccagagccgctcggaactgaatccccaagagaagaagccagatcct 120
Qy 349 gcgccttcctgaaccgagatgctcgccctgcagaagtgcacctaagaagccggaanaa 408
Db 121 gcgccttcctgaaccgagatgctcgccctgcagaagtgcacctaagaagccggaanaa 180
Qy 409 cgggctcgaagagcagatcgagcccatatgaatgaltcaacagaccttgcagagagca 468
Db 181 cgggctcgaagagcagatcgagcccatatgaatgaltcaacagaccttgcagagagca 240
Qy 469 gcgcagagaggtgtggaagagagagtgagtgctctgggaggaagccagatataacagctcc 528
Db 241 gcgcagagaggtgtggaagagagagtgagtgctctgggaggaagccagatataacagctcc 300
Qy 529 agccctcgcgctacaaccgcgcagatcgaggagttatagtcacccggcctgggcttac 588
Db 301 agccctcgcgctacaaccgcgcagatcgaggagttatagtcacccggcctgggcttac 360
Qy 589 taactgtactgtcaggtgcaactttagatgaggggaagagctgtctactaagctgagactg 648
Db 361 taactgtactgtcaggtgcaactttagatgaggggaagagctgtctactaagctgagactg 420
Qy 649 ctggtgagatgtgtgtgtgctgctcgctgagaggaattctcaagcaactgagcgcaagt 708
Db 421 ctggtgagatgtgtgtgtgctgctcgctgagaggaattctcaagcaactgagcgcaagt 480
Qy 709 tccctcgggcccagctcgcgctcgcaggtgtctgagctgtgtgacctgagcgcaagg 768
Db 481 tccctcgggcccagctcgcgctcgcaggtgtctgagctgtgtgacctgagcgcaagg 540
Qy 769 tccctcgtgagatcgcagacccctcccttggccaatctcaagagctgccccctctccacc 828
Db 541 tccctcgtgagatcgcagacccctcccttggccaatctcaagagctgccccctctccacc 600
Qy 829 tacttcgagacttccaggttcaactgagagggcctcggtctccccaagatgtccagagct 888
Db 601 tacttcgagacttccaggttcaactgagagggcctcggtctccccaagatgtccagagct 660
Qy 889 gcgcgtccctcgcagacactctctgagccacgggtccctctgccccaccctcagcgct 948
Db 661 gcgcgtccctcgcagacactctctgagccacgggtccctctgccccaccctcagcgct 720
Qy 949 ctgtgtccagacactgccccctccctctagagagctgctgggagctgttcaagtgatttcca 1008
Db 721 ctgtgtccagacactgccccctccctctagagagctgctgggagctgttcaagtgatttcca 780
Qy 1009 tcccaataataacagtatctcccaactctatctataacactccccaacgcccact 1063
Db 781 tcccaataataacagtatctcccaactctatctataacactccccaacgcccact 835

RESULT 7
AAS03964 standard: DNA: 898 BP.
AAS03964;
26-SEP-2001 (first entry)
Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.
TWEAK extracellular domain: tumour necrosis factor; TNF; angiogenesis; ocular neovascularisation; diabetic retinopathy; neovascular glaucoma; retinoblastoma; retinopathy of prematurity; retrolental fibroplasia; rubecosis; uveitis; macular degeneration; arthritis; rheumatism; ds; corneal graft neovascularisation; psoriasis; metastatic condition; malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint; preneoplastic condition; myocardial angiogenesis; wound granulation; scleroderma; vascular adhesion; telangiectasia; ischaemia; human; atherosclerotic plaque neovascularisation; coronary atherosclerosis; peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR; fusion protein.
Homo sapiens.
Synthetic.
Key Location/Qualifiers
CDS 52..873
/*tag= a
/product= "Fusion protein comprising a growth hormone leader, a leucine zipper multimerisation domain, and human TWEAK extracellular domain"
WO200145730-A2.
28-JUN-2001.
19-DEC-2000; 2000WO-US34755.
20-DEC-1999; 99US-0172878.
10-MAY-2000; 2000US-0203347.
(IMMV) IMMUNEX CORP.
Wiley SR;
WPI: 2001-417975/44.
P-PSDB; AAU03459.
Modulating angiogenesis in a mammal for treating diseases mediated by angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or peripheral tissue, by administering antagonist or agonist of TWEAK receptor -
Example 1; Page 39-40; 46pp: English.
The sequence represents a DNA from the expression vector pDC409-LZ-TWEAK, which encodes a fusion protein comprising a growth hormone leader, a leucine zipper multimerisation domain, and the extracellular domain of human TWEAK. The fusion protein was used in the isolation of human TWEAK receptor (TWEAKR)-expressing clones from a COS cell human cDNA library. The TWEAK protein is a member of the tumour necrosis factor (TNF) family and induces angiogenesis. TWEAKR may therefore be used to screen for and develop TWEAKR agonists and antagonists for the modulation of angiogenesis, to be used in the treatment and diagnosis of human disease. The disorders mediated by angiogenesis include ocular disorders characterised by ocular neovascularisation such as diabetic retinopathy, neovascular glaucoma, retinoblastoma, retinopathy of prematurity, retrolental fibroplasia, rubecosis, uveitis, macular degeneration and corneal graft neovascularisation, and inflammatory diseases such as arthritis, rheumatism and psoriasis. Other treatable diseases include malignant and metastatic conditions such as sarcomas and carcinomas, benign tumours and

CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
CC scleroderma, vascular adhesions, atherosclerotic plaque
CC neovascularisation, telangiectasia, wound granulation, coronary
CC atherosclerosis, peripheral atherosclerosis and ischaemia.
SQ Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other:
Query Match 45.8%; Score 629.2; DB 22; Length 898;
Best Local Similarity 99.5%; Pred. No. 1e-124;
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 232 agttgggagccgggcatcgtctccgccaagagctgcccagagagctgttgca 291
|||||
DB 250 agttgggagccgggcatcgtctccgccaagagctgcccagagagctgttgca 309
QY 292 gaggaggaccagaccgcgcgaactgaatccccaagaaagaaagccagatctcgc 351
|||||
DB 310 gaggaggaccagaccgcgcgaactgaatccccaagaaagaaagccagatctcgc 369
QY 352 ccttcctgaaccgaactagtcgctccgcaagaatgcaactaaaggccgaaacacgg 411
|||||
DB 370 ccttcctgaaccgaactagtcgctccgcaagaatgcaactaaaggccgaaacacgg 429
QY 412 gctcgaagagcgatcgcagcccatatgaagtcatccagacactggaagagcg 471
|||||
DB 430 gctcgaagagcgatcgcagcccatatgaagtcatccagacactggaagagcg 489
QY 472 caggcaggtgtggaagcagtgagtgctggagggaagccagaatacaagctccagc 531
|||||
DB 490 caggcaggtgtggaagcagtgagtgctggagggaagccagaatacaagctccagc 549
QY 532 ccttcgctacaaccgcagacatcgggagttatagtaaccgggcttggtactaac 591
|||||
DB 550 ccttcgctacaaccgcagacatcgggagttatagtaaccgggcttggtactaac 609
QY 592 ctgtactcgaagtgcactttgataagggaagctgttaccctgaagctgacttctg 651
|||||
DB 610 ctgtactcgaagtgcactttgataagggaagctgttaccctgaagctgacttctg 669
QY 652 gtggatggtgtgctgcccctgcgctggagggaatttcagaccagcgccagcttc 711
|||||
DB 670 gtggatggtgtgctgcccctgcgctggagggaatttcagaccagcgccagcttc 729
QY 712 ctcgagccccaagctccgctctgcagagtgcttgaggctgttgccctgcgagcgctcc 771
|||||
DB 730 ctcgagccccaagctccgctctgcagagtgcttgaggctgttgccctgcgagcgctcc 789
QY 772 tccctcgagatccgaaccctccctggggcccatctcaaggtgcgcccttctcaactac 831
|||||
DB 790 tccctcgagatccgaaccctccctggggcccatctcaaggtgcgcccttctcaactac 849
QY 832 ttcggaactcttcaggttaccaggaaggcgccctgg 865
|||||
DB 850 ttcggaactcttcaggttaccaggaaggcgccctgg 883
RESULT 8
ID AAV18599 standard: CDNA: 1168 BP.
AAV18599;
21-JUL-1998 (first entry)
Mus musculus tumour necrosis factor related ligand (TRELL) gene.
TRELL: tumour necrosis factor related ligand; tnfr; treatment;
KW cancer; autoimmune disease; immune system; stimulation; suppression;
CC graft rejection; ds.
OS Mus musculus.
XX

| | | | |
|-------------|--|--------------|--|
| FH | Key | | Location/Qualifiers |
| FT | CDS | | .2..679
/*tag= a |
| ET | | | /note= "tumour necrosis factor related ligand" |
| PN | | MO805783-AI. | |
| PD | | 12-FEB-1998. | |
| XX | | | |
| PF | | 07-AUG-1997; | 97WO-USJ3945. |
| PR | | 18-MAR-1997; | 97US-UO40820. |
| PR | | 07-AUG-1996; | 96US-U0023541. |
| PR | | 18-OCT-1996; | 96US-U0028515. |
| PA | (BIOJ) BIOGEN INC.
(UNGE-) UNIV GENEVA FACULTY MEDICINE. | | |
| PI | Browning JL, Chicheportriche Y; | | |
| PT | Tumour necrosis factor related ligand - useful for, e.g. treating cancer, auto-immune disease and immune responses to tissue grafts | | |
| PS | Claim 2; Pages 45-46; 69pp; English. | | |
| CC | The sequence is that encoding mouse tumour necrosis factor related ligand (TRELL). TRELL or active fragments can be included with a carrier in pharmaceutical compositions to treat cancer, autoimmune diseases or immune responses to tissue grafts, or to stimulate or suppress the immune system. It is useful to screen for TRELL receptors, by labelling with a detectable label and screening compositions for binding. Agents interfering with TRELL-receptor binding can also be screened for, can then be administered, optionally with interferon-gamma, to induce cell death or treat, suppress or alter immune responses (especially involving human adenocarcinoma cells) involving a signal pathway between TRELL and its receptor. The DNA sequence can be used in gene therapy for TRELL-related disorders in mammals (especially humans), e.g. tumours, autoimmune and inflammatory diseases or inherited genetic disorders, by introducing into cells, and expressing, therapeutically effective amounts of a vector, e.g. a virus comprising a gene encoding TRELL. It may also be of use in the preparation of prepare probes for screening natural/synthetic DNAs for TRELL-encoding sequences and for antisense therapy. | | |
| Sequence | 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other: | | |
| Query Match | 44.8%; Score 614.6; DB 19; Length 1168;
Best Local Similarity 75.8%; Pred. No.1.4e-121; | | |
| Matches | 906; Conservative 0; Mismatches 219; Indels 70; Gaps | | |
| Oy | 178 ggcctgggaccttgagcgttcgacgtcgccctgccgccttctctgcgcgtgatcaattg 237

Db 2 gtgcgaagtgcgtgactgctgcgtgccttcgtctcttgtctgtctgctgccccg 61

Oy 238 ggaggcggcatcgatctgtccgcaccagaagcctgtcccagaagaaatggtgcagaagag 297

Db 62 gggagctggacaacgctgtctgccacagagcctttcacgagagagactgtcacagagac 121

Oy 298 gagcagggccgctgcgaactgaatccccaagaaaagaaagcagaatctcgcgtttc 357

Db 122 cgcggggagcccccgtgaactgaatccccaagaaaagaaagcagaatctgtglaaccttc 181

Oy 358 ctgaaccagctaagtctgcgcctgcacgaagtgcaccataaacgacggaanaaacagggctcga 417

Db 162 ttggtaacaactagctgcgcctcgaagaagtgtctcttaaagcgccgaagagcggcgctgcg 241

Oy 418 agagcactgcagcccatatgaattcaatccacagacactygacagagcggcgcagacca 477
 | | |

| | | | |
|-------------------------------------|------|--|------|
| Db | 242 | cgagactatcgagcccatltagagttcattccctcgcgcagagaaagatgtagacaaga | 301 |
| Oy | 478 | ggtgtgagccggagacgtgagttgctctggagaggaagcagaatctcaagctccctctg | 537 |
| Db | 302 | ggtgtgagttgagacagtgtgagttgctgggaagagaccacaatcaacagctccagccctctg | 361 |
| Oy | 538 | cgctacaacgcgcagatcgcggagattatagtaaccgggcctggtctctactactgtac | 597 |
| Db | 362 | cgctacgcgcgcagatctggagaaattatcaacgactcatcagagctctggtctactactgtac | 421 |
| Oy | 598 | tgtcagagtgacactttagatgaagggaagagctgtctactctgaagcttgagctgtgtgat | 657 |
| Db | 422 | tgtcagagtgacactttagatgaagggaagagctgtctactctgaagcttgagctgtgtgat | 481 |
| Oy | 658 | ggctgtgcgcgcgcctctgcctctgcctgaagaaattctcagccactctgcgcagcttccctcgg | 717 |
| Db | 482 | gggtgtgcgcgcgcctctgcctctgcctgaagaaattctcagccacagcagcagctctcctg | 541 |
| Oy | 718 | ccccagctcgcgcctctgcgcaggtgtctgtggctgttgccctgcgcagagtgctcctctg | 777 |
| Db | 542 | ccccagctcgcgttgtgtgcagagtgctgtggctgttgccctgcgcagagtgctcctctg | 601 |
| Oy | 778 | cgagatccgcgcacccctccctctggcccatcttaagagctgcgcgcctctctcaactctgga | 837 |
| Db | 602 | cggaatccgcacccctccctctgggcacttaagagctgcgcgcctctctcaactcttggga | 661 |
| Oy | 838 | ctcttcacaggtctcaatgtagagggccctgtgtctcccaacagtgctccagagctgcgcgtcc | 897 |
| Db | 662 | ctcttcacaggtctcaatgtagagggccctgtgtctcccaatctcttaaaccttccctgtccc | 721 |
| Oy | 898 | -----cctcgacagctctctgtgggacccggttccctctgcgcacccatcagcgtctct | 952 |
| Db | 722 | aggagcatcacacaactccctccacccacccacatccctccacccctc--gctgtcctct | 780 |
| Oy | 953 | gctcagaacgtgcgcctccctctctagaggtcgtctgggcctgttcaagtggtttccatccc | 1012 |
| Db | 701 | ggtccagtcctcgtctctcc--tcaaaaggcagcagagctgtttcacatgtttccatccc | 837 |
| Oy | 1013 | acataatacagtatctcccaactcttaacttaacaactccccacgcgcacactctccacctc | 1072 |
| Db | 838 | -----acagacgtatcctctgtcctcttctaacaatcccatccacacacaactatccaactc | 891 |
| Oy | 1073 | actagctccccaatccctgaacctttagagcccacagtgatctcgactccccctctggcca | 1132 |
| Db | 892 | actagctccccaagccctac-----tlatccctgactccccacccact | 936 |
| Oy | 1133 | cagagccccagagcattgtgtctcaactgtactctgtgtggcaagatgtgtccagaagacc | 1192 |
| Db | 937 | caccgcagcaagtggtttattgtacttctgtgac----- | 968 |
| Oy | 1193 | cacttcagagcaactaaggggcctgagacctgcgcagaggaagccaagagactgggctag | 1252 |
| Db | 969 | -----caggcaactgtagtggtgcctgagaccgtgtgtgcagaggaagccaagaaacccgtggactag | 1022 |
| Oy | 1253 | gcacgaagttcccaaatgtgtaggggcgagaaacaa--gacaagctccctcccttgaagattc | 1311 |
| Db | 1024 | gcacgaagttcccaacgtgtgtaggggcgagagcgtggggacaagctcctccctgta----tc | 1079 |
| Oy | 1312 | cctgtgtgatttttaaacacatatatttttatattatattgtgacaaatgttga | 1366 |
| Db | 1080 | cctgtgtgatttttgaaa--agatactatttttatattatattgtgacaaatgttga | 1132 |
| RESULT 9 | | | |
| AAAX23425 | | | |
| ID AAAX23425 standard; DNA; 701 BP. | | | |
| XX AAAX23425: | | | |
| XX | | | |
| XX 18-JUN-1999 (first entry) | | | |
| XX | | | |
| DE Mouse TNRL3 DNA. | | | |
| XX | | | |

(HYSEQ INC.

PA Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
XX Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lanson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Rita D, Garcia V, Jones LW, Stache-Grain B;
XX
DR WPI: 2000-126369/11.
XX
PT polynucleotide library used to determine cancerous states of mammalian
XX cells -
XX
PS Claim 1; Page 994; 1097pp; English.
XX
CC AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed
CC genes correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived.
CC The polynucleotides sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of
CC pre-metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer.
XX
SQ Sequence 1000 BP; 21 A; 53 C; 574 G; 22 T; 330 other;

Query Match 4.4%; Score 60.6; DB 21; Length 1000;
Best Local Similarity 36.1%; Fred. No. 0.0012;
Matches 171; Conservative 0; Mismatches 303; Indels 0; Gaps 0;

OY 668 cccttgcgctgcgtgaggaattctcagcaacgtgcggacgttcctcgtggcccaagtccc 727
DB 911 CCGGGCCCCCCCCCGGNCNCCCCTCCCNNNCCCGANNCCCGCAGCCGCCNCCNCC 852
OY 728 gcctctgcagaagtgtctgggtgttctggcccttgaggcagagttcctccctgcatccga 787
DB 851 CCCGNNNNNNNCCCGGGGNGCCGGCGGNGNCCCGCCNCCCGCCCTTCCGCCGCC 792
OY 788 cccttccccgtggccatctcaaggtgtcccttccctaacttctggaattcttcagg 847
DB 791 CCCGNNGGAGNCCCNMTCCCGCCNCCNCCCGNCCCGCCNCCNCCCGCCCGCCNT 732
OY 848 tttaactgaggagccttgttccccaagtgctccaaggtctgcggtccctcctgaagc 907
DB 731 CCCCCCNCNCGNCCNCCNNGNCCNCCCGCCCGCCCGCCGCGNCCNCCCGCCNCCNT 672
OY 908 tctctggagcacccagttccctctgcgccaccctcagacgcgttcttgctcagacctgcc 967
DB 671 CGCGNCCNCCNCCNCCGTTCNCCGTTCNCCCGCCCGCCCGCCCGCCNCCNCCNCCNCC 612
OY 968 ctccctctagaaggtgtcgtggcctgttcaagtgltttccataccaacaataacagtat 1027
DB 611 CNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 552
OY 1028 tccaccttatctataaacatcccccaacagccaaacttcccaacttaactgtctcccaalc 1087
DB 551 NCCCGCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 492
OY 1088 ctgtagaccttgaggaccccaagtgatctcgaactcccccttgccacagagcccc 1141
DB 491 CCNCGNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 438

RESULT 14

[illegible]

```
QY 422 cgatcgagaccatgatgaatgcatcaccagaccctgagcagagcagcgagcagtg 481
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10085 C-TTDSNSTRCYASTYDASRRCSYSTYRASRGTBSDDCTBSRSCYSYDASRYANCDTB 10027
QY 482 tggagcggagcagtgatgctgctggagaggaagcagaatcaaacgctccagccctcgct 541
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10026 CYTTBAKTRACYDAVYARCNSTSYRAYSDSTCYTRCGRSNCSTYSYSTYASRTTB 9967
QY 542 acaaccgccaatcgaggagcttatagaccaggcgctggcctctaccctgctatctc 601
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9966 TTTCYCTBCRCRCRCYSSTCNYSCTCYTSRSTTTSTCTCTTSTNSTBTBYS 9907
QY 602 aggtgacacttgatgagggaggaagcgtctacactgaaactggaactgctggtgagtg 661
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9906 YSSTYS-----SRGYSRSGSDSRGNCYNTSTNCYDASTDFTBYSRCYTYVSYSIDS 9855
QY 662 tgcggcgccctgctgctgaggaatctcagccacgcggcgccagcttcctcgggccc 721
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9854 TDSSTDSSTSTBDSCTYTTTBSRSTSDSTSTYKCSRDYDATTBDSNSTNCCYDASRTBTB 9795
QY 722 agctccgctctgccaagtgctcggtgctgctgagcccgccaaggtccctccctgcgga 781
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9794 STNCTARKYTBTDARCSDSSTYSRGIYDANSTSRYSSTYSSTYSYSAKYCAKSTTB 9735
QY 782 tccgacccctccctcgggccatctcaagctgcgccctctccctcactt-cggaact 840
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9734 TBCYDADACDYADANCSYSSDSTYBYCSRCCYDAYASCSRYDARCYDACSYSNS 9675
QY 841 ttccaggtctcagtgagggccctgctcctcccaagctgctccagcgcgcgcgcctccct 900
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9674 TCYDATTBTTSRYSTTYSYSNCTYDATTSTRCBYSTBTBTSCAKCTBDSSTANSTST 9615
QY 901 cgacagcctctgagcagcgcgtc-cctctcgcccacccccaagcgcgcctctgtccag 959
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9614 YSTRCTRYSSRSRCSYSCSRSCSRSCYSDSTCYSTYTYTAASCTTSRSTSDA 9555
QY 960 acctgcctccctctcagagcgtgcctgagcgtctcagtgcttccatccacataaa 1019
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9554 SRSSTYSRCTTSTYSSTYSDSDSTSTBTBNTYSSDSDCTBYSRSDRSDSTCNC 9495
QY 1020 taagatctccactctatcttaactccccccagccacactcctcactcaactaact 1079
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9494 YSCSDSRSTTYDACYTDAAKTBCTYTSNDCCNSTSRCTYSCSRTSRSTCCCT 9435
QY 1080 cccaatccctgaccccttgaggccccaagtgatctgaccccccctcgccagcagacc 1139
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9434 BTSTGNCCTDAYDANSTSRDAYDACYYSVDASTTBYSCTBYSSTYDAYSRSYC 9375
1140 c 1140
Db 9374 Y 9374
```

```
PD 18-NOV-1999.
XX
XX 13-MAY-1999; 99WO-US10602.
XX
XX 14-MAY-1998; 98US-0085426.
XX 15-MAY-1998; 98US-0085537.
XX 15-MAY-1998; 98US-0085636.
XX 21-OCT-1998; 98US-0105234.
XX 27-OCT-1998; 98US-0105877.
XX
XX (CHIR ) CHIRON CORP.
XX (HISE-) HISEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
XX Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
XX Planson G, Dmanac R, Crivenjakov R, Dickson M, Dmanac S, Labat I;
XX Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
XX WPI; 2000-126369/11.
XX
XX Polynucleotide library used to determine cancerous states of mammalian
XX cells -
XX
XX Claim 1; Page 995-996; 1097pp; English.
XX
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
XX libraries constructed from human colon cancer cell lines. The present
XX invention also describes a method of detecting differentially expressed
XX genes correlated with a cancerous state of a mammalian cell, comprising
XX detecting at least one differentially expressed gene product in a test
XX sample derived from a cell suspected of being cancerous, where detection
XX of the differentially expressed gene product is correlated with a
XX cancerous state of the cell from which the test sample was derived.
XX The polynucleotide sequences can be used in a method for detecting
XX differentially expressed genes correlated with a cancerous state of a
XX mammalian cell. The polynucleotides can also be used as probes for
XX detecting and mapping related genes. They can be used in diagnosis and
XX prognosis of diseases and disorders (e.g. identification of
XX pre-metastatic or metastatic cancerous states, stages of cancer, or
XX responsiveness of cancer to therapy). This is particularly for breast
XX cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
XX negative breast cancer, lung cancer, and colon cancer.
XX
XX Sequence 1218 BP; 9 A; 31 C; 494 G; 37 T; 647 other;
XX
XX Query Match 4.2%; Score 58; DB 21; Length 1218;
XX Best local similarity 27.1%; Pred. No. 0.0044;
XX Matches 118; Conservative 0; Mismatches 318; Indels 0; Gaps 0;
XX
XX QY 704 cgaattccctcgagccagctcgccctctgccaagtgtctggtgctggtgacctgagc 763
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1107 CCCNNNNNNNCCCCCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1048
QY 764 cagggtctctccctcgagatccgacactccctcggtggcccatcgaagctgcccccttc 823
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1047 CCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 988
QY 824 tcacctactcgtgactctcaggtcaggtcaggggacctggtctccacagctgctcc 883
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 987 NNCNCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 928
QY 884 agctgctcggtctccctcgacagctctctgagcagccggtccctctgcccacactcag 943
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 927 NNCGCCGCCGCCGCCGCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 868
QY 944 ccgctcttgctcagactgccccctccctctaaaggtcgtgctggtcgttcaagtgtc 1003
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 867 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 808
QY 1004 ttccatccacataataacagatattccactctatacttaaacctcccccaagccact 1063
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 807 TCCNNNNNNCCCCCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 748
```


THIS PAGE BLANK (USPTO)

264 LeuProTPrAlaHisLeuLysAlaAlaProPheLeuThrTyPheGlyLe 280
|||||
601 CTCCTGCGGGCCATCTCAAGCTGC.CCCTTCTCAGCTACTTCGAGCT 649
280 uPheGlnValHis 284
|||||
650 CTCACAGGTTCCAC 662
seq_name: gb_est2:BI819200
seq_documentation_block: 777 bp mRNA linear EST 04-OCT-2001
LOCUS BI819200 603034614F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175698 5',
DEFINITION mRNA sequence.
ACCESSION BI819200
VERSION BI819200.1 GI:15930750
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 777)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11437 row: 1 column: 03
High quality sequence stop: 759.
Location/Qualifiers
1. 777
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5175698"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 140 a 268 c 215 g 154 t
ORIGIN
alignment_scores: Length: 188
Quality: 979.00 Gaps: 0
Ratio: 5.207
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-245-198a-4 x BI819200 ..
Align seg 1/1 to: BI819200 from: 1 to: 777
97 AlaGluGlnAspGlnAspProSerGluLeuAsnProGlnThrGluGln 113
|||||
14 GCACAGAGGAGCACGACCGCTCGAAGTGAATCCCGACAGAGAAGAAAG 63
|||||
113 rGlnAspProAlaProPheLeuAsnArgLeuValArgProArgArgSerA 130

|||||
64 CCAGGATCTCGCGCTTCTCTGAACCGCACTGATCGGCTCGCAGAAAGT 113
|||||
130 laProLysGlyArgLysThrArgAlaArgArgAlaIleAlaAlaHisTyr 146
|||||
114 CACCTAAAGGCCGGAACACAGGCTCGAAGAGGATCGCAGCCCATAT 163
|||||
147 GluValHisProArgProGlyGlnAspGlyValGlnAlaGlnValAspGly 163
|||||
164 GAAGTTCTATCCACGACCTCGACAGGACGAGCGACGAGGAGGTGTGACG 213
|||||
163 yThrValSerGlyTrpGluGlnAlaArgIleAsnSerSerProLeuA 180
|||||
214 GACAGTGAAGTGCTGGAGGAGAAACAGATCAACAGCTCCAGCCCTTGC 263
|||||
180 rGlyTrsAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr 196
|||||
264 GCTACAAACCCGACAGATCGGGAGTTATATGTCACCCGGGCTGTAC 313
|||||
197 TyrLeuTyrcysGlnValHisPheAspGlyGlyLysAlaValTyrLeu 213
|||||
314 TACCTGTACTGTCAAGTGCATTGTATGAGGGAAGGCTGTCTACTTGA 363
|||||
213 sleuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGln 230
|||||
364 GCTGACTTGTCTGGATGGATGTGTGCTGCCCTCGCTGCGTGAAGAA 413
|||||
230 heserAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgCysGln 246
|||||
414 TCTCAGCCACGCGCGCGCTCCCTCGGCGCCAGCTCCGCTCTGCAC 463
|||||
247 ValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgTyr 263
|||||
464 GTGCTGTGGGCTGTGGCTCGGCGCAGGCTCTCCCTCGGATCCGAC 513
|||||
263 rLeuProTPrAlaHisLeuLysAlaAlaProPheLeuThrTyPheGly 280
|||||
514 CTCGCCCTGGGCGCATCTCAAGGCTGCCCTTCTCTCACTTCTGAC 563
|||||
280 eupheGlnValHis 284
|||||
564 TCTTCCAGGTTCCAC 577
seq_name: gb_est2:BI766766
seq_documentation_block: 834 bp mRNA linear EST 25-SEP-2001
LOCUS BI766766 603056866F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5',
DEFINITION mRNA sequence.
ACCESSION BI766766
VERSION BI766766.1 GI:15758344
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 834)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11517 row: c column: 18
High quality sequence stop: 772.
Location/Qualifiers
1. 834
FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5206217"
/clone_1db="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site_1: NOT; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb. Insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT      154 a      287 c      223 g      170 t
ORIGIN

```

```

Alignment_block:
  Alignment: 962.00      Length: 194
  Ratio: 5.010          Gaps: 1
  Percent Similarity: 98.969      Percent Identity: 98.969

```

US-09-245-198a-4 x B1766766 ..

Align seg 1/1 to: B1766766 from: 1 to: 834

```

92 GINGLUGLULeuValAAGLUGLUSerGlnAspProSerGluLeuAsnPr 108
   |||||||
   3 CAGGAGAGCTGTGTGCGAGAGGAGCCAGGAGCCCTCGAAGTGAATCC 52
108 oGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeuValA 125
   |||||||
53 CCAGACAGAGAAAGCCAGAGATCTGCGCTTCTCGAAGCCGACTGATTC 102
125 rGProArGArSerAlaProLysGlyArgLysThrArgAlaArgArgAla 141
   |||||||
103 GGCCTCGCAGAGTGCACCTAAAGCCGGGAAACAGGGGCTCGAAGAGCG 152
142 TLeAlaAlaHIStyrGluValAlaHISProArGPro.GlyGlnAspGlyAlaG 158
   |||||||
153 ATCGCAGCCCATTAATGACATTCACACGACCGTGCAGAGGAGGAGCGC 202
158 TAlaGlyValAlaSpGlyThrValSerGlyTTPGluGluAlaArgIleAsn 174
   |||||||
203 AGGCAAGTGTGACGAGGACAGTGAAGTGGCTGGGAGAGAACCCAGAAATCAAC 252
175 SerSerSerProLeuArGtyrAsnArGlnIleGlyGluPheIleValTh 191
   |||||||
253 AGCTCAGCCCTGCGGTACAAACCCGAGATCGGGAGCTTATTAAGTACAC 302
191 rArgAlaGlyLeuTyrTyrLeuTyrCysGlnValAlaHISPheAspGluGly 208
   |||||||
303 CCGGCGTGGGCTTACTACTGATGTCAGGTCGACTTGTAGAGAGGGA 352
208 yAlaValAlaTyrLeuLysLeuAspLeuLeuValAspGlyValLeuAlaLeu 224
   |||||||
353 AGGCTTCTACTGAGAGCTGGACTTCTGGTGGATGTGTGCTGGCCCTTG 402
225 ArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProG 241
   |||||||
403 CGCTGCTGGAGAAATTCACGCACTGGGCGCAATTCCTCGGGGCCCA 452
241 nLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArGProGlySers 258
   |||||||
453 GCTCCGCTCTGCGCAGGTGTGGGCTGTGGCCCTGCGGCCAGGTGC. 501
258 rLeuLeuArgIleArGtyrLeuProTyrAlaAlaHISLeuLysAlaAlaProPhe 274
   |||||||
502 CCCTGGGATCCGACACCTCCCTGGGGCCCATCTCAAGGCTGCCCTTTC 551

```

```

275 LeuThrTyrPheGlyLeuPheGlnValHis 284
   |||||||
552 CTCACCTACTCGAGCTTTCAGGTTTCAC 581

```

seq_name: gb_est2:B1596681

seq_documentation_block:

```

LOCUS      B1596681      828 bp      mRNA      linear      EST 07-SEP-2001
DEFINITION 603243254F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892 5',
mRNA sequence.
ACCESSION  B1596681
VERSION    B1596681.1 GI:15489620
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE 1 (bases 1 to 828)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

Tissue Procurement: Mikhailos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: L16M11722 row: k column: 13

High quality sequence stop: 776.

Location/Qualifiers

1..828

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5285892"

/clone_1db="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptPR (modified

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag

); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',

size-selected for average insert size 2.3 kb and

normalized to R0T 5. This is a primary library enriched

for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIH/NHGRI, National

Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 155 a 278 c 223 g 172 t

ORIGIN

Alignment_scores:

Quality: 934.00 Length: 180

Ratio: 5.189 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 99.444

Alignment_block:

US-09-245-198a-4 x B1596681 ..

Align seg 1/1 to: B1596681 from: 1 to: 828

```

105 GILUeLAsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAs 121
   |||||||
44 GAACGTGAATCCCGACAGCAAGAAAGCCAGAGATCCGCGCTTCTGTGA 93
121 nArgLeuValArGProArGSerAlaProLysGlyArgLysThrArgA 138
   |||||||
94 CCGACTAGTTGGGCTCGCAGAAAGTGCACCTAAAGCCGGAAACAGCGG 143
138 TArGArGAlaIleAlaAlaHIStyrGluValAlaHISProArGProGlyGln 154

```


seq_name: gb_est1:AW763237

seq_documentation_block:

LOCUS AW763237 561 bp mRNA linear EST 04-MAY-2000

DEFINITION ur70d09.y1 NCI-CGAP Mam3 Mus musculus cDNA clone IMAGE:1155633 5'

similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS

;/ mRNA sequence.

AW763237

AW763237.1 GI:7695174

EST.

KEYWORDS house mouse.

SOURCE Mus musculus.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 561)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Other-ESTs: ur70d09.x1

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov/image/html/lresources.shtml

MG1:1058389

Seq primer: -40RP from Gdbco

High quality sequence stop: 433.

FEATURES

source

1..561

/organism="Mus musculus"

/strain="129 - C57/Bl6 - EVBN"

/db_xref="taxon:10090"

/clone="IMAGE:3155633"

/clone_lid="NCI CGAP Mam3"

/tissue_type="tumor, gross tissue"

/dev_stage="10 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT 108 a 158 c 194 g 100 t 1 others

ORIGIN

Alignment_scores:

Quality: 825.00 Length: 187

Ratio: 4.609 Gaps: 0

Percent Similarity: 95.722 Percent Identity: 86.096

alignment_block:

US-09-245-198a-4 x AW763237 ..

Align seg 1/1 to: AW763237 from: 1 to: 561

41 SerGlnArgArgArgGlyArgGlyGluProGlyThrAlaLeuLeuA 57

|||||

1 AGCCAGAGCGCGAGGGGCGCGGGGAGCCGCGGACCGCCTGCTGGC 50

57 lProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaLeuAlaLeuL 74

|||||

51 CCGGCTGGTGTAGGCTGCGCTGGCTGGCTGGCTGGCTGGCTGGC 100

74 euAlaValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluPro 90

|||||

101 TGGTCGTGCTACGCTGGGAGCTGGGCAACGCTGCTGCCAGAGCCT 150

91 ALaGlnGluGluLeuValAlaGluGluAspGlnAspProSerGluLeuAs 107

|||||

151 TCTCAGAGAGAGCTACAGACAGACAGACCGCGGAGCCCTGAACTGAA 200

107 nProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeu 124

|||||

201 TCCCGACAGAGAGAAAGCAGAGATGTGTACTCTTCTTGGAACTAG 250

124 aArgProArgArgSerAlaProLysGlyArgGlyThrArgAlaArgArg 140

|||||

251 TCCGGCTCTGAGAGAGTCTCTTAAAGCCGGAGGCGGCGCTGCGGA 300

141 AlaIleAlaAlaHisThrGluValHisProArgProGlyGlnAspGlyAl 157

|||||

301 GGTATTGCAAGCCATTATGAGTTCATCTCGGCCAGACAGAGATGAGAC 350

157 aGlnAlaGlyValAspGlyThrValSerGlyTrpGluGluAlaArgIleA 174

|||||

351 ACAAGCAGGTGTGATGGACAGTGAGTGGGAGAGACCAAAATCA 400

174 snSerSerSerProLeuArgGlyTrpAsnArgGlnIleGlyGluPheVal 190

|||||

401 ACAGCTCCAGCCCTCTGCGCTACGACCCGACATGGGGAATNTACAGTC 450

191 ThrArgAlaGlyLeuThrTyThrLeuTyrcysGlnValHisPheAspGlu 207

|||||

451 ATCAGGCGGCGCTACTACTACTGCTACTGCTGAGTGCACATTGATGAGAG 500

207 yLysAlaValTyThrLysLeuAspLeuValAspGlyValLeuAlaLe 224

|||||

501 AAAGCTGTCTACTGAAAGCTGAGCTGCTGTGAACGCTGTGCGGCC 550

224 euArgCysLeu 227

|||||

551 TCGCTGCGCTG 561

seq_name: gb_est2:BT62908

seq_documentation_block:

LOCUS BT62908 785 bp mRNA linear EST 25-SEP-2001

DEFINITION 603047966F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5188168 5',

mRNA sequence.

ACCESSION BT62908

VERSION BT62908.1 GI:15754486

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 785)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

plate: L14M11470 row: c column: 17

High quality sequence stop: 647.

FEATURES

source

1..785

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5188168"

/clone_lid="NIH_MGC_116"

/lab_host="DH10B"

/note="Organ: pooled colon, kidney, stomach; Vector: PCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: This is a NIH-MGC Library."

BASE COUNT 124 a 228 c 274 g 159 t
ORIGIN

alignment_scores:

Quality: 742.50 Length: 258
Ratio: 3.519 Gaps: 18
Percent Similarity: 81.783 Percent Identity: 75.581

Alignment block:
US-09-245-198a-4 x B1762908 ..

Align seg 1/1 to: B1762908 from: 1 to: 785

```

25 G1YAlaValArgGlnAlaGlnProAlaProMeta1aAlaArgArGS 41
|||||
29 GGGGGTGAAGTCACTGCACAGACCCCGCCCACTTGGCTCCCGTCCGA 78
41 erGlnArGArGArgG1YArGArg.G1Y.GluProGly...Thra1aLeu1 56
|||||
79 GCCAGAGCGGAGGGGGCGGTGTGTGAGCTGGGCACTGGCCCTCT 128
56 euValaProLeuAlaLeuGly...LeuG1YLeuAla1.LeuAlaG1Sleu.. 70
|||||
129 TGTGTCTCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 178
71 .....G1YLeuLeuAlaValSerLeuG1YSerArGAlas 84
|||||
179 GCCCTCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 222
84 erLeuSerAlaGlnGlnProAlaGln.GluGluLeuValaGlu...G1 99
|||||
223 CTGTCTGCGCCAGTGAAGCTGCCAGTGAAGAGCTGTGTGTGTGTGTGT 272
99 uASG1aSPProSer...GluLeuAsnProGlnThrGluGluSerGlnA 115
|||||
273 TGACCACTGACCCCTGTGTGAATCCCAACAGACAAAGAGCCAGG 322
115 sPProAlaProPhelLeuAsnArg..LeuValaRPProArGArGSerAlaP 131
|||||
323 ATCTGCGCCCTTCTCTGACCTGACTACTGTGGCTGCGACAAAGTGACC 372
131 OlySG1YArGlyThrArGAlaArGArGAla1LeAla1aH1sTYrGluV 148
|||||
373 TAAAGCCCGGAACACGCGGCTCGAAGAGCATGCGACCCATTATGAG 422
148 aH1sPProArGProG1YLeuAspG1YAlaGlnAlaG1YAlaSPG1YTh 164
|||||
423 TTTCAATCCAGCACTGTGACAGAGGAGGAGGAGGAGGAGGAGGAGCA 472
165 Val1SerG1YTPGluGluAlaArg1LeuAsnSerSerSerProLeuArGT 181
|||||
473 GTTGAAGTGGCGGAGAGAGCCAGAAATCAACAGCTCCAGCCCTTGGCCT 522
181 YrAsnArGln1LeG1YLeuPhe1LeuValThraGAlaG1YLeuTYrTYr 197
|||||
523 ACAACCGCCAGATCGGGAGTTTATAGTACACCGGCGCTCGTCACTAGTAC 572
198 LeuTYrGlySGlnValH1sPheAspGluG1YrGlyAlaValTYrLeuV 214
|||||
573 CTGTACTGTCTAGTCACTTTGATGAGGAGGAGGAGGAGGAGGAGGAGT 622

```

```

214 uASpLeuValaSPG1YAlaLeuAla...LeuArG1SleuGluGlnP 230
|||||
623 GGACTTGTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 672
230 heSerAlaThAlaAla1.SerSerLeu.G1YProGlnLeuArGLeuCYG 246
|||||
673 TCTCAGGCACTGCGCGGAGTTCCTCGGGGCCAGCCAGCTCCGCTCC 722
246 InValSer.G1YLeuLeu1AlaLeuArGProG1YSerSerLeuArG1eA 262
|||||
723 AGGTGTCTGGGGCTGTGTGGCTTGGGGGAGGAGGAGGAGGAGGAGTCC 772
262 rG 262
773 GG 774

```

seq_name: gb_est2:BG686319

seq_documentation_block:

LOCUS BG686319 587 bp mRNA linear EST 01-MAY-2001
DEFINITION 602638232F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4766071 5',

ACCESION BG686319

VERSION BG686319.1 GI:13917716

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 587)

NIH-MGC Institutes of Health, Mammalian Gene Collection (MGC)

unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

tissue procurement: Louis M. Staudt, M.D., Ph.D.

cDNA library preparation: Ling Hong/Rubin Laboratory

cDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LICM1625 row: P column: 08

High quality sequence stop: 587.

Location/Qualifiers

1..587

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4766071"

/tissue_type="primary B-cells from tonsils (cell line)"

/lab_host="DH10B (phage-resistant)"

/note="Organ: B-cells; Vector: pOTB7; Site.1: XhoI;

Site.2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCAAGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by Ling

Hong in the Laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: This is a NIH-MGC Library."

BASE COUNT 108 a 187 c 205 g 87 t

ORIGIN

alignment_scores:
Quality: 738.00 Length: 156
Ratio: 4.855 Gaps: 2
Percent Similarity: 97.436 Percent Identity: 97.436

alignment_block:
US-09-245-198a-4 x BG686319 ..

gene 1. .433
/gene="Tnfsf12"
CDS <1. .433
/gene="Tnfsf12"
/note="data source:MCD, source key:MGI:1196259,
evidence:ISS
putative
tumor necrosis factor (ligand) superfamily, member 12"
/codon_start=2
/protein_id="BAB32249.1"
/db_xref="GI:12861641"
/translation="IAAHYEVHRRPQDAQAQAVDCTVSGWEETKINSSPLRYDRI
GEPYIRAGILYLYCOVHDEKAVYIKLIDLNVGVIALCLEEFSAATASSPQDLR
LCOVSGLLPRPSSLRIRITLPAHLKAPFLTYFGLFOVH"
BASE COUNT 228 a 300 c 248 g 256 t 1 others
ORIGIN

alignment_scores:
Quality: 707.00 Length: 143
Ratio: 5.086 Gaps: 0
Percent Similarity: 97.203 Percent Identity: 94.406

alignment_block:
US-09-245-198a-4 x AK020909 ..

Align seg 1/1 to: AK020909 from: 1 to: 1033

142 ILEAIAIAHISTYRGLIVAHISPROARGPROGLYINASPGIYALGI 158
|||||
2 ATTCAGAGCCATTATGAGGTTTCATCTCGCCAGACAGATGAGACACA 51
|||||
158 nAAGLYVALASPGIYTHRALSERGLYTRPGIUGLUALARGYLEANS 175
|||||
52 AGCAGGTGTGATGGACAGTGAAGTGGAGAGACCAAAATCAACA 101
|||||
175 erSerSerProLeuArGTYrAsnArGInIleGlyGluPheIleValThr 191
|||||
102 GCTCAGAGCCCTCTGGCTAGACAGCCGCAATGGGGAATTACAGTCATC 151
|||||
192 ArgAlaGlyLeuYrTYrLeuYrCYsGlnValHisPheAspGlyLy 208
|||||
152 AGGGCTGGGCTCTACTACTACTAGTCAGGCTGACTTTGATGAGGAAA 201
|||||
208 sAlaValYrLeuYsLeuAspLeuValAspGlyValLeuAlaLeuA 225
|||||
202 GGCTGTACTCTGAAGCTGAGCTGTGTGAGAGGTGTGCTGGCCCTGC 251
|||||
225 rGcYsLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGln 241
|||||
252 GCTGCTGTGAGAAATTCACAGCCAGCAGCAGACGCTCTCCGGGCCAG 301
|||||
242 LeuArgLeuCYsGlnValSerGlyLeuLeuAlaLeuArgProGlySer 258
|||||
302 CTCGGTTGTGCTCAGAGTGTCTGGGCTGTGGCGCGCCAGGGCTCTC 351
|||||
258 rLeuArgLeuArgThrLeuProTrpAlaHisLeuYsAlaAlaProPheL 275
|||||
352 CTTGGAGTCCGACACCTCCCTGGGCTCATCTTAAGGCTGCCCTTCC 401
|||||
402 TAACCTACTTTGGACTCTTTCAGATTCCAC 430
|||||

seq_name: gb_est1:AW917574

seq_documentation_block:
LOCUS AW917574 584 bp mRNA linear EST 25-MAY-2000
DEFINITION EST348878 Rat gene index, normalized rat, norvegicus, Bento Soares
Rattus norvegicus cDNA clone RGIERF49 5' end, mRNA sequence.
ACCESSION AW917574
VERSION AW917574.1 GI:8083328
KEYWORDS EST.

SOURCE Norway rat.
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 584)
AUTHORS Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat
Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information
Seq primer: M3 Reverse.

FEATURES
source
1. .584
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="RGIERF49"
/clone_lib="Rat gene index, normalized rat, norvegicus,
Bento Soares"
/tissue_type="mix - brain, ovary, placenta, kidney, lung,
liver, embryo, heart, muscle, spleen"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); site_1: EcoRI; site_2:
XhoI; Estimated insert size approx 1 kb"
BASE COUNT 108 a 188 c 147 g 138 t 3 others
ORIGIN

alignment_scores:
Quality: 652.00 Length: 135
Ratio: 5.015 Gaps: 0
Percent Similarity: 96.296 Percent Identity: 92.593

alignment_block:
US-09-245-198a-4 x AW917574 ..

Align seg 1/1 to: AW917574 from: 1 to: 584

150 ProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSe 166
|||||
2 CCACAGCCAGACAGAGATGAGACAGCAGGTGTGATGGAGACAGTGAG 51
|||||
166 rGlyTrpGluGluAlaArgIleAsnSerSerSerProLeuArgYrAsnA 183
|||||
52 TGGCTGGGGAAGAGACCAAAATCAACAGCTCCAGCCCTGTGGCTATGAC 101
|||||
183 rGlnIleGlyGluPheIleValThrArgAlaGlyLeuYrTYrLeuYr 199
|||||
102 GCCAGATTGGGGAATTTACGGTCATCAGGCGCTGACTACTACTGTAC 151
|||||
200 CysGlnValHisPheAspGluGlyLysAlaValYrLeuYsLeuAspLe 216
|||||
152 TGTCAAGTGCACTTTGATGAGGGGAGGAGGAGCTGCTGAAGCTGACT 201
|||||
216 uLeuValAspGlyValLeuAlaLeuArgCYsLeuGluGluPheSerAla 233
|||||
202 GCTGTGATGATGTGTGTGGCTGTGCTGCTGCTGGAATAATTCACAGCA 251
|||||
233 hTrAlaAlaSerSerLeuGlyProGlnLeuArgLeuCYsGlnValSerGly 249
|||||
252 CAGACGACAGCTCTCTCCGGCCAGAGCTCCGTGTGCTGACAGTGTGG 301
|||||
250 LeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTr 266
|||||
302 CTGTTGCTCTGTGGGCCAGGGTCTTCCCTCGAGATCCCTACCTCCCTCG 351
|||||

```
266 pAlaHisLeuLysAlaAlaProPheLeuThyTyrPheGlyLeuPheGlnY 283
|||||
352 GGCCTATCTTAAAGGCTGCCCCCTTCTCTACTTACTTGACCTCTNCAAG 401
|||||
283 aAlHis 284
|||||
402 TTCAC 406

seq_name: gb_est2:BI906850

seq_documentation_block:
LOCUS BI906850 698 bp mRNA linear EST 16-OCT-2001
DEFINITION 603064633F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5213480 5',
mRNA sequence.
ACCESSION BI906850
VERSION BI906850.1 GI:16169619
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 698)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1536 row: b column: 09
High quality sequence stop: 613.
Location/Qualifiers
1..698
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5213480"
/clone_lib="NIH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV
(destroyed); RNA source: leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

BASE COUNT 149 a 204 c 243 g 102 t
ORIGIN

alignment_scores:
Quality: 648.50 Length: 150
Ratio: 4.665 Gaps: 3
Percent Similarity: 92.667 Percent Identity: 92.000

alignment_block:
US-09-245-198a-4 x BI906850 ..
Align seg 1/1 to: BI906850 from: 1 to: 698

16 ProArSerLeuGlySerArGAspGlyAlaValArGlnAlaGlnPr 32
|||||
15 CCCCAGATCTCGGTCGCCGGGATGGGGCGGTGAGGACGACACGCC 64
|||||
32 oProAlaProMetAlaAlaArGArSerGlnArGArGlyArGArGc 49
```

```
|||||
65 CCCCCGCCCCATGCGCGCGCTCGAGCCAGAGCGGAGGGGCGCGCG 114
|||||
49 lYgluProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeu 65
|||||
115 GGGAGCCGGGACACCGCTGTGTGTCCCTGCGCTGGGCTGGGCTG 164
|||||
66 AlAlaLeuAlaCysLeuGlyLeuLeuAlaValAlaSerLeuGlySerAr 82
|||||
165 GCTGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 214
|||||
82 g.AlAserLeuSerAla...GlnGluProAlaGlnGluGluLeuAla 97
|||||
215 GGGCATCGCTGCTGCCGCCAGAGACCTGCCAGAGAGAGCTGTGGCA 264
|||||
98 GluGluAspGlnAspProSerGluLeuAsnProGlnThGluGluSerG 114
|||||
265 GAGGAGGACGAGACCGCTGGAGACTGAATCCCGACAGAGAAAGAGCA 314
|||||
114 nAspProAlaProPheLeuAsnArGLeuValArGProArGArSerAlaP 131
|||||
315 GATCTCTGGCGCTTCTCTGACCGACTAGTTCGCTGCAGAGATGCAC 364
|||||
131 rOlYsGlyArGlySThrArGAlaArGArAlaAlaAlaHisTyrGlu 147
|||||
365 CTAAAGCCCGGAAMACACGCGCTCGAAGAGCATGCCACCATTAATGA 414
|||||
148 ValHisProArGProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 163
|||||
415 GTTATCTCAAGACTGACAGAGGAGGAGCGGAGCA.....GATGGA 456

seq_name: gb_est2:BI908274

seq_documentation_block:
LOCUS BI908274 894 bp mRNA linear EST 16-OCT-2001
DEFINITION 603068526F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5217367 5',
mRNA sequence.
ACCESSION BI908274
VERSION BI908274.1 GI:16171193
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 894)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1546 row: d column: 08
High quality sequence start: 5
High quality sequence stop: 460.
Location/Qualifiers
1..894
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5217367"
/clone_lib="NIH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV
(destroyed); RNA source: leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
```


1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH-MGC Library."

BASE COUNT 244 a 260 c 253 g 137 t
ORIGIN

alignment_scores:

Quality: 647.00 Length: 204
Ratio: 3.615 Gaps: 16
Percent Similarity: 87.745 Percent Identity: 82.843

alignment_block:
US-09-245-198a-4 x B1908274 ..

Align seg 1/1 to: B1908274 from: 1 to: 894

```

29 GlnAlaIProProAlaProMetaAlaAlaArgSer...GlnArgAr 44
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
3 CAGGACATGCCCCCCCCCATTCCTGCTGATGCCAGAGCG 52
44 gArg. GlyArgArgGly...GlnProGlyThraAlaLeuValPro.L 59
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
53 GATGTGGGGCGCCCTGGGGTGATGCCGGGACACTGCTGTCCTGC 102
59 euAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeu.GlyLeuLeuVal 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
103 TCGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGC 152
75 aValValSerLeuGlySerArgAlaSerLeuSer.AlaGlnIProAla 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
153 CGTGTGCTGATTTGGGGAGCCGGGATCCCTCTCCGCCAGAGACTGCC 202
92 GlnGlnIuLeuValAla.GlnIuAspGlnAspProSerGlnLeuAsp 108
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
203 CAGGAGAGAGCTGCGCATGAGAGACAGACCCGCGAAGTGAATC 252
108 roGlnIuGlnIuSerGlnAspProAlaProPheLeuAsnArgLeuVal 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
253 CCCGAGAGAGAAAGCCAGATCCCTGCTTCTGAACCGACTAGTT 302
125 ArgProArgSerAlaProLysGlyArgLysThraArgAlaArgArgAl 141
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
303 CGGCTGCGAGAAAGTGCACCTAAAGCCGGAACACGCGCTCGAAGAGC 352
141 aAlaAlaAlaHisTyrGluValHisProArgProGlnAspGlyAlaG 158
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
353 GATCGAGAGCCATATGAAATTATCCAGACTGAGACAGAGAGAGCGCC 402
158 InaIaGlyValAspGlyThrValSer.GlyTyrPheGlnIuAlaArgIleAs 174
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
403 AGGAGAGTGTGAGAGGAGAGCTGCGTGGAGAGAAAGCCAGAAATCAA 452
174 nSerSer.SerProLeu.ArgTyrAsnArgGlnIleGlyIuPhe.Ile. 189
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
453 CAGTCTCAGAGCCCTGAGCGCTACACCGACAGATCGGGAGCGTAAATAG 502
190 ValThrArgAlaGlyLeu.TyrTyrLeuTyrCys.GlnVal.HisPheAs 205
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
503 GTCACCGGGGCTGGGCTGACTACTGCTACTGACAGTGCACACTTTGA 552
205 pGlu.GlyLysAlaValTyrLeuLysLeuAspLeu 216
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
553 TGAGCGGGAAGGCTGCTACTACTGAAGACTGAACCTT 587

```

seq_name: gb_est2:B1965174

seq_documentation_block:

LOCUS B1965174 474 bp mRNA linear EST 23-OCT-2001
DEFINITION id34e07.y1 Melton Normalized Mixed Mouse Pancreas 1 NI-MSI Mus
musculus cDNA 5' similar to TR-043508 043508 TNF-RELATED WEAK
INDICER OF APOPTOSIS ; mRNA sequence.

ACCESSION B1965174
VERSION B1965174.1 GI:16339579
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Maira,M., Pape,D., Wylie,T., Martin,J., Blistein,A., Schmitt,A., Theising,B., Rittler,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvill,I., R., Williams,T., Jackson,Y., and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other_ESTs: id34e07.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biochem.harvard.edu

LIBRARY was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
High quality sequence stop: 453.
Location/Qualifiers
1..474
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1 NI-MSI"
/sex="Both for embryonic & newborn, male for adult and adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn, adult, mixed"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site 1: Not I; Site 2: Sal I; Five libraries representing E10.5/12.5 pancreatic bud, E16.5 pancreas, newborn pancreas, adult pancreas, and adult islets of Langerhans were separately constructed using Superscript Plasmid Library kit (Life Technologies). cDNA was made by oligo-dT priming and size-selected by column fractionation. Libraries were amplified once on solid support and plasmid DNA from each library was prepared and mixed in equal amounts. The mixed library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded mixed library plasmid DNA was mixed with 5 micrograms PCR product representing mixed library inserts and hybridized to an Eco of 6. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

FEATURES
source

alignment_scores:
Quality: 645.50 Length: 155
Ratio: 4.421 Gaps: 1
Percent Similarity: 94.194 Percent Identity: 85.806

alignment_block:
US-09-245-198a-4 x B1965174 ..

Align seg 1/1 to: B1965174 from: 1 to: 474

14 ProLeuProArgSerLeuGlySerArgAspGlyAlaValArgGlnAl 30

```

13 CCTCCCA...TCCCTCGGGTCCCGGGATGGGGGGCGGTGAGGAGCGC 59
30 aglInProAlaProMetAlaAlaArgArgSerGlnArgArgGlyA 47
60 ACAGCCCCCGCCCATGGCCCGCTGGAGCCAGAGCGAGGGGGC 109
47 rGArGlyGluProGlyThrAlaLeuValProLeuAlaLeuGlyLeu 63
110 GCGGGGGGAGCGGGGACCGCCCTGCTGGCCCGCTGGTGTGAGCGTG 159
64 GtyleuAlaLeuAlaCysleuGlyLeuLeuAlaValValSerleuG 80
160 GCGCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 209
80 ySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluValA 97
210 GAGCTG.TTTATGTTGTCTGCCCCAGAGGCTTCTCAGAGAGCTGAC 258
97 laGluGluAspGlnAspProSerGluLeuAsnProGlnThrGluSer 113
259 CAGAGGACCGCGGGAGCGCCCTGAACCTGAATCCCGACAGAGGAAC 308
114 GlnAspProAlaProPheLeuAsnArgLeuValArgProArgSerAl 130
309 CAGGATGTGCTACCTTCTTGGACAACTAGTCCGCTCGAAGAGTGC 358
130 aProLySGlyArgLySThrArgAlaArgArgAlaIleAlaIleHisT 147
359 TCCTAAAGCCGGAAGCGCGCTGGCCGAGCTATTGCGAGCCCATTA 408
147 lValAlaHisProArgProGlyGlnAspGlyAlaGlnAlaGlyVal 163
409 AGGTTCATCTCGCGCAGACAGAGATGAGCAACAGAGGTGTGGAG 458
164 ThrValSerGlyTrp 168
459 ACAGTGAAGTGGCTGG 473

seq_name: gb_est1:AA870722

seq.documentatn_block: 445 bp mRNA linear EST 16-MAR-1998
LOCUS AA870722
DEFINITION vq25967.r1 Barstead stromal cell line MRLRB8 Mus musculus cDNA
ACCESSION AA870722
VERSION AA870722.1 GI:2966167
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 445)
Matta,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucada,T., Lacy,M., Le,M., Martin,D., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:601556
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 420.
Location/Qualifiers

```

```

source
1. .445
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1095324"
/cell_line="C2C12 (undifferentiated)"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site:1: EcoRI; Site:2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TCCTACCAATCTGACAGCGAGCGCGCCCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGAATCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Source undifferentiated tissue culture cell line C2C12.
Library constructed by Bob Barstead. The C2C12 cell line
(available from ATCC, catalog # CRL-1772) differentiates
rapidly, forming contractile myotubes and producing
characteristic muscle proteins."

BASE COUNT 102 a 119 c 135 g 89 t
ORIGIN

alignment_scores:
Quality: 604.00 Length: 138
Ratio: 4.576 Gaps: 0
Percent Similarity: 95.652 Percent Identity: 86.957

alignment_block:
US-09-245-198a-4 x AA870722 ..

Align seg 1/1 to: AA870722 from: 1 to: 445
109 GlnThrGluSerGlnAspProAlaProPheLeuAsnArgLeuValAr 125
12 CAGGCTGCGAAGGCCAGAGATGTGCTTCTTGGACAACTAGTCCG 61
125 pProArgArgSerAlaProLySGlyArgLySThrArgAlaArgArgAlaI 142
62 GCCTGCAAGAGTGCCTTAAAGCGCGGAGAGG. GCGCTCGCGAGCTA 110
142 lValAlaHisTyrGluValAlaHisProArgProGlyGlnAspGlyAlaGln 158
111 TTGCACGCCCATTAATGAGTTCATCCCGGCCAGAGACAGATGAGCAAA 160
159 AlaGlyValAspGlyThrValSerGlyTrpGluGluAlaArgGlnLeu 175
161 GCAGGTGTGAGATGGACAGTGAAGTGGTGGGAAGACCAAAATCAACAG 210
175 rSerSerProLeuArgTyrAsnArgGlnIleGlyIupheIleValThra 192
211 CTCACACCTCTGCGCTACGACCGCCAGATGGGGAATTTACAGTCATCA 260
192 rGAlaLyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlyGly 208
261 GGGCTGGGCTTACTACTGTACTGTACGTGACACTTGTATGAGGGAAG 310
209 AlaValTyrLeuLysLeuAspLeuValAspGlyValLeuAlaLeuAr 225
311 GCTGTACTCTGAAGCTGAGCTTGTGTGAACGGGTGTCTGGCCCTGGC 360
225 gCysLeuGluGluPheSerAlaThrAlaAlaSerLeuGlyProGlnL 242
361 CTCGCTGGAAGAAATTTCTACGCCACACAGCAAGCTCTCTGGGCCCCAC 410
242 euaArgLeuCysGln 246
411 TCGGTTGTGGCCAG 424

```

Mon Jun 24 10:53:41 2002

OM of: US-09-245-198a-4 to: N_Geneseq_032802:* out_format: pfs

Date: Jun 22, 2002 2:32 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=frame+r2n_model -DEV=xlh
-O=/cgn2.1/USPTO_spool/US09245198/runat_21062002.184008.9995/app-query.fasta_1.632
-DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-EGAPOP=6.000 -FGAPEXT=7.000 -YGAPEXT=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=biosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09245198 @CGN1_1_0 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1
```

Search information block:

```
Query: US-09-245-198a-4
Query length: 284
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 287.880000
```

score_list:

| Sequence | Strd Orig | ZScore | Escore | len | Documentation |
|---|-----------|---------|---------|-----|---------------|
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV18600 + | 1444.00 | 1871.03 | 7.2e-96 | 13 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV56000 + | 1377.00 | 1783.63 | 5.3e-91 | 14 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV49717 + | 1366.00 | 1769.76 | 3.1e-90 | 13 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV47613 + | 1265.00 | 1639.24 | 5.9e-83 | 12 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV43350 + | 1265.00 | 1639.24 | 5.9e-83 | 12 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV43350 + | 1066.00 | 1383.33 | 1.1e-68 | 8 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV23424 + | 1062.00 | 1376.94 | 2.4e-68 | 10 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV18599 + | 1020.00 | 1321.24 | 3.0e-65 | 11 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV23425 + | 968.00 | 1258.09 | 1.0e-61 | 70 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV24140 + | 127.00 | 152.58 | 3.76 | 28 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV49205 + | 126.50 | 156.65 | 2.23 | 16 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV69289 + | 124.50 | 149.40 | 5.66 | 28 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAV25519 + | 124.50 | 144.56 | 4.93.20 | 15 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV09686 + | 124.00 | 145.84 | 8.93 | 3 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAV25519 + | 123.50 | 113.91 | 536.08 | 15 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV62160 + | 123.50 | 135.74 | 32.60 | 11 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV62160 + | 123.50 | 115.68 | 427.27 | 11 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV46674 + | 120.50 | 144.21 | 11.00 | 28 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV42224 + | 117.50 | 137.31 | 25.99 | 35 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV79598 + | 117.50 | 137.51 | 25.99 | 39 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAV70738 + | 117.50 | 137.51 | 26.00 | 39 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV99230 + | 117.00 | 148.46 | 6.38 | 10 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV25514 + | 117.00 | 148.46 | 6.38 | 10 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAV21473 + | 116.00 | 147.61 | 7.12 | 9 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:AAV01620 + | 115.00 | 144.00 | 11.31 | 12 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAV01318 + | 115.00 | 143.54 | 11.99 | 13 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:AAV03062 + | 115.00 | 143.53 | 12.00 | 13 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV4455 + | 115.00 | 142.61 | 13.51 | 1 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV21793 + | 115.00 | 133.60 | 42.94 | 4 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV41035 + | 115.00 | 131.14 | 58.82 | 5 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV2729 + | 115.00 | 127.97 | 88.34 | 6 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV62133 + | 114.50 | 123.33 | 160.14 | 12 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAV05397 + | 113.50 | 132.46 | 49.69 | 38 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAV78508 + | 113.50 | 111.14 | 765.10 | 44 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAV78508 + | 113.50 | 111.14 | 765.10 | 44 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAV29605 + | 113.00 | 131.20 | 58.37 | 41 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAV29604 + | 113.00 | 130.20 | 66.31 | 46 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAV73500 + | 113.00 | 124.94 | 130.26 | 84 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV63350 + | 113.00 | 116.93 | 364.16 | 21 | |
| /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAV63348 + | 113.00 | 107.42 | 1.2e+03 | 63 | |

```
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV6029 + 112.50 140.19 18.42
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:AAV51445 + 112.50 136.66 29.00
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAV29515 + 112.50 132.26 51.00
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAV43352 + 112.50 131.91 53.32
/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAV4843 + 112.50 130.99 59.99
```

seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV18600

seq_documentation_block:

ID AAV18600 standard; cDNA; 1373 BP.

AAV18600;

21-JUL-1998 (first entry)

Homo sapiens tumour necrosis factor related ligand (TRELL) gene.

TRELL; tumour necrosis factor related ligand; tnfr; treatment;

cancer; autoimmune disease; immune system; stimulation; suppression;

graft rejection; ds.

Homo sapiens.

key Location/Qualifiers

CDS 1..852

18-MAR-1997; 970S-0040820.

07-AUG-1996; 960S-0023541.

18-OCT-1996; 960S-0028515.

(BIOI) BIOGEN INC.

(OYGE-) UNIV GENEVA FACULTY MEDICINE.

Browning JL, Chlcheportliche Y;

WPI: 1998-145619/13.

P-PSDB; AAV47525.

Tumour necrosis factor related ligand - useful for, e.g. treating

cancer, auto-immune disease and immune responses to tissue grafts

Claim 2: Pages 48-50; 69pp; English.

The sequence is that encoding human tumour necrosis factor related

ligand (TRELL). TRELL or active fragments can be included with a

carrier in pharmaceutical compositions to treat cancer, autoimmune

diseases or immune responses to tissue grafts, or to stimulate or

suppress the immune system. It is useful to screen for TRELL

receptors, by labelling with a detectable label and screening

compositions for binding. Agents interfering with TRELL-receptor

binding can also be screened for, can then be administered,

optionally with interferon- gamma, to induce cell death or

treat, suppress or alter immune responses (especially involving human

adenocarcinoma cells) involving a signal pathway between TRELL and its

receptor. The DNA sequence can be used in gene therapy for

TRELL-related disorders in mammals (especially humans), e.g. tumours,

autoimmune and inflammatory diseases or inherited genetic disorders,

by introducing into cells, and expressing, therapeutically effective

amounts of a vector, e.g. a virus comprising a gene encoding TRELL.

It may also be of use in the preparation of prepare probes for

screening natural/synthetic DNAs for TRELL-encoding sequences

and for antisense therapy.

Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other;

alignment_scores:
 Quality: 1444.00 Length: 284
 Ratio: 5.085 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-245-198a-4 x AAV18600 ..

Align seg 1/1 to: AAV18600 from: 1 to: 1373

```

1 MetSerLeuLeuAspPheGluIleSerAlaArgLeuProLeuProAr 17
|||||
1 ATGTCATTGTTAGACTTTGAATTTCCGCCGCCGCTCCCTCCCTCCCG 50
17 gSerLeuGlySerArgAspGlyAlaValAlaValArgIleAlaInProProA 34
|||||
51 ATCCCTCGGGTCCGGGATGGGGGGGGGTGAGCGAGCGACAGCCCGCG 100
34 laProMetAlaAlaArgArgSerGlnArgArgArgGlyArgArgGlyGlu 50
|||||
101 CCCCATGGCCGCCGCTCGAGCAGAGCGGAGGGGGCGCGGGGGAGAG 150
51 ProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuAlaLe 67
|||||
151 CCGGGCACCGCCCTGCTGCTCCGCTCGCGCTGGCGCTGGCGCTG 200
67 uAlaGlySerLeuGlyLeuLeuLeuAlaValAlaSerLeuGlySerArgAla 84
|||||
201 GCCCTGCTCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 250
84 erLeuSerAlaGlnGluProAlaGlnGluGluLeuValAlaGluGluAsp 100
|||||
251 CGCTGCGCGCCAGAGAGCTGCCAGAGAGAGCTGCTGCTGCTGCTG 300
101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAl 117
|||||
301 CAGGACCGCTCGGAATGATCCAGACAGAAAGCAAGATCCTGCTG 350
117 aProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyAla 134
|||||
351 GCCTTTCTCGAACCGACTAGTTGCGGCTCGCAGAGTGCACCTTAAGGCC 400
134 rGlySerThrArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisPro 150
|||||
401 GGAAGAACAGGGCTCGAAGAGCGATCGCAGCCATTATGATTCATCCA 450
151 ArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerG 167
|||||
451 CGACCTGGACAGAGCGAGCGAGCGAGGTGTGAGCGGAGCAGTAGTGG 500
167 yTrpGluGluAlaArgGlnLeaSerSerSerSerProLeuArgTyrAsnArg 184
|||||
501 CAGGGAGGAGCCAGATCAACAGCTCCAGCCCTGCTGCTGCTCAACCGCC 550
184 InIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrG 200
|||||
551 ACATCGGGAGTTTATAGTCACCCGGCTGGGCTGCTACTACTGTACTGT 600
201 GlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeu 217
|||||
601 CAGGGAGCTTTGATGAGGGGAGAGGCTGCTACCTGAACTGGACTTGTCT 650
217 uAlaIleAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrA 234
|||||
651 GGTGGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 700
234 laAlaSerSerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeu 250
|||||
701 CGGCAGATTCTCGGGGCCCAAGCTCGGCTCTGCGCAGGTGTCTGGGCTG 750
251 LeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTrrAl 267
|||||

```

```

751 TTGGCCCTGCGGCCAGGGTCTCTCCCTGCGGATCCGACCTCCCTGGGCG 800
267 ahIleuLysAlaIleProPheLeuThrTyrPheGlyLeuPheGlnValH 284
|||||
801 CCATCTCAAGGCTGCCCTTCTCTCCTACTACTGCGACTTTCAGGTTTC 850
284 Is 284
||
851 AC 852

seq_name: /SIS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAx56000
seq_documentation_block:
ID AAX56000 standard; DNA: 1421 BP.
XX
XX
AC AAX56000:
XX
DF 15-JUL-1999 (first entry)
XX
DE Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.
XX
KW Human: tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
KW cancer; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 92..841
FT /*tag= a
FT /*product= "Apo-3 ligand"
XX
XX MO9919490-A1.
XX
XX PD 22-APR-1999.
XX
XX PF 09-OCT-1998; 98WO-US21407.
XX
XX PR 17-DEC-1997; 97US-0069862.
XX PR 10-OCT-1997; 97US-0062037.
XX
XX PA (GETH ) GENENTECH INC.
XX
XX PI Ashkenazi AJ, Marsters SA, Pittl R;
XX WP1: 1999-287982/24.
XX DR P-PSDB: AAY0369.
XX
XX PS New human Apo-3 ligand (a tumour necrosis factor) homologue
XX
XX Claim 18; Fig 1; 74pp; English.
XX
CC The present sequence encodes a human tumour necrosis factor (TNF) and
CC lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
CC cytosolic activity. Apo-3 ligand can be used to induce apoptosis in
CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and
CC to induce JNK/SAPK-dependent responses in mammalian cells.
XX
XX SO Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other;

alignment_scores:
Quality: 1377.00 Length: 278
Ratio: 5.062 Gaps: 0
Percent Similarity: 97.842 Percent Identity: 97.842

alignment_block:
US-09-245-198a-4 x AAX56000 ..

Align seg 1/1 to: AAX56000 from: 1 to: 1421

7 GluIleSerAlaArgArgLeuProLeuProArgSerLeuGlySerArgAs 23
|||||

```

5 GAGATCCCTCGACCTGCAGCCACGCGTCCGCGATCCTCGGCTCCGGGA 54
xx
23 pGLyAlaValaArgInaIagInProProAlaProMetaIaIaIaArga 40
|||||
55 TGGGGGGGGGTGAGGACAGCAGCCGCCGCCCATGAGCCGCCGCTC 104
40 rGserInaArgArgArgIaArgArgIaIagInProIaIaIaIaIaIaIa 56
|||||
105 GGAGCCAGAGGGGAGGGGGGGGGGGAGCCGGGAGCCGCCGCTGCTG 154
57 ValProIeuaIaIeuaIyIeuaIyIeuaIaIeuaIaIyIeuaIyIe 73
|||||
155 GTCCCGCTCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGCT 204
73 uIeuaIaValaIaSerIeuaIyIeuaIyIeuaIaIaSerIeuaIaIagIn 90
|||||
205 GCTGGCGGTGCTCAGTTGGGGAGCCGGGCTGCTGCTGCTGCTGCTG 254
90 roaIaIagInIagInIaIeuaIaIagInIaIaIaIaIaIaIaIaIa 106
|||||
255 CTGCCAGAGAGAGCTGTGGCAGAGAGAGAGAGAGAGAGAGAGAGAG 304
107 AspProIa 123
|||||
305 AATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354
123 uValaIa 140
|||||
355 AGTTGGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 404
140 rGaIa 156
|||||
405 GAGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 454
157 AlaIa 173
|||||
455 GCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 504
173 eAsnSerSerSerProIeuaIaIaIaIaIaIaIaIaIaIaIaIaIa 190
|||||
505 CAACAGCTCCAGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 554
190 AlaIa 206
|||||
555 TCACCCGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
207 GLyIa 223
|||||
605 GGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
223 AlaIa 240
|||||
655 CTGGCGCTGCTGAGAGATTCAGCCACTGCGGCGAGATTCCTCGGGC 704
240 roIa 256
|||||
705 CCCAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 754
257 SerSerIeuaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 273
|||||
755 TCCCTGCTGGAGATCCAGCCCTGCTGCTGCTGCTGCTGCTGCTG 804
273 oPhaIeuaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 284
|||||
805 CTTCCTCACTACTTCGAGCTCTTCAGAGTTTAC 838
seq_name: /STD1/9c9data/geneseq/geneseq-emb1/NA2000.DAT:AAA49717

seq_documentation_block:
ID AAA49717 standard: cDNA: 1353 BP.
AC
XX
XX AAA49717:
XX

DT 25-SEP-2000 (first entry)
XX
DE Human PRO207 cDNA clone DNA30879-1152.
XX
KW PRO207; human; antitumour; tumour; therapy; cytostatic;
KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;
KW uterine cancer; prostate cancer; lung cancer; bladder cancer;
KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 58..807 /*tag= a
FT sig_peptide 58..177 /*tag= b
FT mat_peptide 178..804 /*tag= c
FT
XX
PN WO200037638-A2.
XX
PD 29-JUN-2000.
XX
PF 02-DEC-1999; 99WO-US28565.
XX
PR 22-DEC-1998; 98US-0113296.
PR 08-MAR-1999; 99WO-US05028.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA:
PI Napier MA, Pitti RM, Wood WI;
XX
XX WPI: 2000-442668/38.
DR P-PSDB: AAY95338.
XX
XX Novel composition to inhibit neoplastic cell growth or for treating
PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or
PT PRO866
XX
XX Claim 20: Fig 3; 172pp; English.
XX
PS The present sequence is that of cDNA clone DNA30879-1152
CC (ATCC 209358) encoding human PRO207 (see AAY95338), which shows
CC homology to several members of the tumour necrosis factor family,
CC especially human lymphotoxin (23.4%). The cDNA was identified in a
CC foetal kidney cDNA library following identification of an expressed
CC sequence tag with homology to human Apo-2 ligand. A claimed method
CC for inhibiting the growth of a tumour cell comprises exposing the
CC tumor cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,
CC PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866 (see
CC AAY95337-49), their agonists or chimeric polypeptides incorporating
CC them. The tumour is especially a cancer selected from breast,
CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and
CC central nervous system cancer, melanoma and leukaemia. Nucleic
CC acids encoding PRO179 etc. are used in the recombinant production
CC of the antitumour polypeptides.
XX
SQ Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

alignment_scores:
Quality: 1366.00 Length: 268
Ratio: 5.097 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-245-198A-4 x AAA49717 ..

Align seg 1/1 to: AAA49717 from: 1 to: 1353

```

17 ArgSerLeuGlySerArgAspGlyAlaValArgGlnAlaGlnProPr 33
   |||
   |||
   |||
1 CGATCCGTCGCGGATGGGGGCGGTGAGGACAGACACACCC 50
33 cAlaProMetAlaAlaArgArgSerGlnArgArgArgGlyArgArgGly 50
   |||
   |||
   |||
51 CCCCCCAAGGCCCGCGGAGCCAGAGCGGAGGGGCGCCCGGGGG 100
50 lProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAla 66
   |||
   |||
   |||
101 AGCCGGGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 150
67 LeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAl 83
   |||
   |||
   |||
151 CTGGCCTGCGCGCTCCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 200
83 aSerLeuSerAlaGlnGlnProAlaGlnGlnLeuValAlaGlnGln 100
   |||
   |||
   |||
201 ATCGCTGTCGCCCGGAGAGCCCTGCCCAGAGAGAGCTGTGGCAGAGAG 250
100 sPGlnAspProSerGlnLeuAsnProGlnThrGlnGlnSerGlnAspPro 116
   |||
   |||
   |||
251 ACCAGAGACCCGTCGGAACGTGAATCCCGACAGAGAAAGCCAGATCCT 300
117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysG 133
   |||
   |||
   |||
301 GCGCCTTTCCTGAACCGACTACTGCGCTCCAGAGAGTGAAGTAAAG 350
133 yArgIysThrArgAlaArgArgAlaIleAlaIleHisThrGlnValHis 150
   |||
   |||
   |||
351 CCGGAAACACGGGCTCGAAGAGCCGATCCACCCCATTTTGAAGTTCAAC 400
150 rOArPProGlyGlnAspGlyAlaGlnAlaGlnValAspGlyThrValSer 166
   |||
   |||
   |||
401 CACGACCTGGACAGGAGCGGAGCGGAGGTGTGGACGGGACAGAT 450
167 GlyTPGlnGlnAlaArgIleAsnSerSerProLeuArgTyrAsnAr 183
   |||
   |||
   |||
451 GCGTGGAGAGCCAGATCAACAGCTCCAGCCCTGCTACACACCG 500
183 gGlnIleGlyGlnPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrC 200
   |||
   |||
   |||
501 CCAGATCGGGAGTTATAGTCACCCGGGCTGCGCTCTACTACT 550
200 ySGlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeu 216
   |||
   |||
   |||
551 GTCAGGTGCACTTTCATGAGGGGAGCGCTGCTACTGTAAGTGGACTTG 600
217 LeuValAspGlyValLeuAlaLeuArgCysLeuGlnGlnPheSerAlaTh 233
   |||
   |||
   |||
601 CTGGTGGAGAGGTGTGCTGCGCTGCGCTGAGGAATTCACAGCCAC 650
233 rAlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyL 250
   |||
   |||
   |||
651 TCGGGGAGATTCCTCGGGCCCAAGCTCCGCTGTCACAGTGTCTGGGC 700
250 euLeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTrp 266
   |||
   |||
   |||
701 TGTGTGCGCTTGGCGGCGAGGTCCTCCCTCGGATCCGACACCTCCCTG 750
267 AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVa 283
   |||
   |||
   |||
751 GCCCATCTCAAGGCTGCCCTTCTCACCCTACTTGAGACTCTCCAGGT 800
283 His 284
   |||
   |||
801 TCAC 804

```

seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV47613

seq_documentation_block:

ID AAV47613 standard; cDNA; 1236 BP.

XX AAV47613;

DT 27-OCT-1998 (first entry)

DE TNF related endothelium proliferative agent gene.

KW ss; TNF; endothelium proliferative agent; TREPA; wound healing; cancer;

KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..750

FT /*lag= a

FT /*product= "TREPA"

PN M09835061-A2.

PD 13-AUG-1998.

PF 12-FEB-1998; 98MO-US02859.

PR 10-FEB-1998; 98US-0021706.

PR 12-FEB-1997; 97US-0798692.

PA (ABBO) ABBOTT LAB.

PI Willey SR.

DR WPI: 1998-447255/38.

DR P-PSDB: AAM29745.

PS Claim 11; Page 123-4; 142pp; English.

The TNF-related endothelium proliferative agent (TREPA), or its active forms or agonists, are used to treat a deficit of TREPA, e.g. to promote wound healing or tissue grafting, by promoting vascularisation, CC also to induce apoptosis for treating cancer and eliminating autoreactive T cells, as an adjunct to cancer chemotherapy or antiviral treatment. CC TREPA peptides can also be used to target cytotoxic agents or for CC affinity isolation of the corresponding receptor, the nucleic acid for CC which can be used to transform tumour cells to render them more CC responsive to TREPA and to screen for TREPA mimics. CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting CC vascularisation), inflammation or a wide range of autoimmune conditions, CC conditions involving abnormal stimulation of epithelial cells (e.g. CC atherosclerosis), for birth control (inhibiting ovulation and placental CC formation) or other angiogenic conditions (e.g. ulcers).

XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

alignment_scores:

Quality: 1265.00

Ratio: 5.080

Percent Similarity: 100.000

Percent Identity: 99.598

alignment_block:

US-09-245-198A-4 x AAV47613 ..

Align seg 1/1 to: AAV47613 from: 1 to: 1236

36 MetAlaAlaArgArgSerGlnArgArgGlyArgArgGlyGluProG 52

1 ATGGCGCCGCTGGAGACCAGAAAGGAGGGGGCCGGGGGAGACCGGG 50
52 ythralaLeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaC 69
51 CACCGCCCTCTGTCGCGCTCGCGCTGGGCTTGGGCTGGCCCTGGCCT 100
69 yslLeuGlyLeuLeuAlaValValSerLeuGlySerArgAlaSerLeu 85
101 GCGTCGGCCCTCTGCTGCTGGCGTGCTCAgTTTGGGAGCCGGGCTGCTG 150
86 SerAlaGlnLeuProAlaGlnGlnLeuValAlaGlnLeuAspGlnAs 102
151 TCGGCCCGAGAGCCTGGCCAGAGAGACTGTTGGCAAGAGAGACAGGA 200
102 proSerGlyLeuAsnProGlnThrGlnGlnSerGlnAspProAlaPro 119
201 CCGCTGGAACTAAATCCAGACAGAAAGCAAGCAGATCTGCGCTT 250
119 heLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgLys 135
251 TCGTGAACCGACTAGTTCGGGCTCGAAGAAAGTGCACCTAAAGCGGAAA 300
136 ThrArgAlaArgArgAlaAlaAlaAlaHisTyrGlnValHisProArgPr 152
301 ACACGGGCTCGAAGAGCGATCGGACGCCCATTTAGAAATTGATCCAGACC 350
152 ogLysAspGlyAlaGlnAlaGlyAlaAspGlyThrValSerGlyTyrPg 169
351 TGGACAGAGAGAGAGCGGACGAGGTGGAGCGGACAGTGAAGTGGCTGG 400
169 IuGlnIuAlaArgIleAsnSerSerProLeuArgTyrAsnArgGlnIle 185
401 AGGAAGCCGAATCAACAGCTCCAGCCCTTGCGCTACAACCGGCAATC 450
186 GlyClnPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnVa 202
451 GGGGAGTTTATGATCAACCGGGGCTGGGCTGTACTACTGTACTGTCAAGT 500
202 IHisPheAspGlnGlyLysAlaValTyrLeuLysLysAspLeuValAla 219
501 GCACCTTTGATGAGGGGAGGCTGTCTACCTGAAGCTGGACTGTGCTGGT 550
219 spGlyValLeuAlaLeuArgCysLeuGlnGlnIuPheSerAlaThrAlaAla 235
551 ATGGGTGGCTGGCCCTGCGCTCGCTGGAGAAATTTCAAGCACTGGCGGG 600
236 SerSerLeuGlyProGlnLeuAlaGlyLeuCysGlnValSerGlyLeuAla 252
601 AGTTCCCTGGGGCCCAAGCTCCGCTGTGGCAGGTGTCTGGGCTGTTGGC 650
252 AleuArgProGlySerSerLeuArgIleArgThrLeuProTPrAlaHis 269
651 CCGCGGCGCCAGGGTCTCCCTCGGGGAATCCACACCTCCCTCGGGGCCATC 700
269 LeuYslAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284
701 TCAAGGCTGGCCCTTCTCACTACTTGGAGACTTTTCAGAGTTTCAC 747

| | | |
|----|--|---|
| XX | Homo sapiens. | |
| OS | | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | CDS | 1..750 |
| FT | | /*tag= a |
| FT | | /product= "Human TREPA (TNF related endothelium |
| FT | | proliferative agent)." |
| XX | | |
| PN | US6207642-B1. | |
| XX | | |
| PD | 27-MAR-2001. | |
| XX | | |
| PF | 26-JUN-1998; | 98US-0105343. |
| XX | | |
| PR | 12-FEB-1997; | 97US-0798692. |
| PR | 10-FEB-1998; | 98US-0021706. |
| XX | | |
| PA | (ABBO) ABBOTT LAB. | |
| XX | | |
| PI | Wiley SR; | |
| XX | | |
| DR | WPI: 2001-280760/29. | |
| XX | | |
| DR | P-PSDB: AAE00891. | |
| XX | | |
| PT | Inducing angiogenesis in mammal at desired sites for promoting wound healing, by administering soluble fragment of extracellular domain of tumor necrosis factor related endothelium proliferative agent protein | |
| XX | - | |
| XX | Example 2; Column 73-74; 53pp; English. | |

```

119 heleAsnArgLeuValArgProArgSerAlaProLysGlyArgLys 135
    |||||||
251 TCCTGAACCGACTAGTGGCGCTCGAAGACGACCTAAAGCGCGGAAA 300
    |||||||
136 ThrArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgPr 152
    |||||||
301 ACAGGGGCTCGAAGAGCGATCGACGCCATTATGAGTTCAATCCAGACC 350
    |||||||
152 oGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPG 169
    |||||||
351 TGGACAGAGGAGGAGCGACGAGGTGTGACGGGACAGTGAAGTGGCTGG 400
    |||||||
169 IuGluValArgIleAsnSerSerProLeuArgTyrAsnArgGlnIle 185
    |||||||
401 AGGAAGCCAGATCAACAGCTCCAGCCCTGCGGTACACCGCCAGATC 450
    |||||||
186 GlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnVa 202
    |||||||
451 GGGGAGTTTATAGTCACCCGGGCTGGCTTACTACTGTCTACAGGT 500
    |||||||
202 IHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuVala 219
    |||||||
501 GCACCTTATGAGGGAGAGGCTGTCTACCTGAAGCTGAGCTGGTGG 550
    |||||||
219 SPGLYValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235
    |||||||
551 ATGGTGCTGCTGGCCCTGGCTGCGCCGAGGAATTCACGCCACCTCGGCG 600
    |||||||
236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAl 252
    |||||||
601 AGTTCCCTCGGGGCCCTCCGCTCCGCCAGGTGTCTGGGCTGTGGC 650
    |||||||
252 AlenArgProGlySerSerLeuArgIleArgThrLeuProTrpAlaHisL 269
    |||||||
651 CCTGGGGCGAGGGTCTCTCCCTGGGATCCGACCTCCCTCGGGCCATC 700
    |||||||
269 euLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284
    |||||||
701 TCAAGGCTGCCCCCTTCTCAACCTACTCTCGGACCTCTCCAGGTTTCAC 747

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: AAS03964
seq_documentation_block:
ID AAS03964 standard; DNA: 898 BP.
XX
AC AAS03964:

26-SEP-2001 (first entry)

Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.
XX
TWEAK extracellular domain: tumour necrosis factor; TNF; angiogenesis;
XX ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
XX retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
XX rubecosis; uveitis; macular degeneration; arthritis; Rheumatism; ds;
XX corneal graft neovascularisation; psoriasis; metastatic condition;
XX malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
XX preneoplastic condition; myocardial angiogenesis; wound granulation;
XX scleroderma; vascular adhesion; telangiectasis; ischaemia; human;
XX atherosclerotic plaque neovascularisation; coronary atherosclerosis;
XX peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;
XX fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
XX
key Location/Qualifiers
FH 52..873
FT CDS /*tag= a
FT /product= "Fusion protein comprising a growth hormone
FT leader, a leucine zipper multimerisation
FT domain, and human TWEAK extracellular

```

```

FT domain"
XX
PN MO200145730-A2.
XX
PD 28-JUN-2001.
XX
PF 19-DEC-2000; 2000MO-US34755.
XX
PR 20-DEC-1999; 99US-0172878.
XX 10-MAY-2000; 2000US-0203347.
XX (IMMV ) IMMUNEX CORP.
XX
XX Willey SR:
XX WPI: 2001-417975/44.
XX DR P-PSDB; AAU03459.
XX
XX Modulating angiogenesis in a mammal for treating diseases mediated by
XX angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
XX peripheral tissue, by administering antagonist or agonist of TWEAK
XX receptor
XX
XX Example 1: Page 39-40; 46pp: English.
XX
XX The sequence represents a DNA from the expression vector
XX pDC409-LZ-TWEAK, which encodes a fusion protein comprising a growth
XX hormone leader, a leucine zipper multimerisation domain, and the
XX extracellular domain of human TWEAK. The fusion protein was used in
XX the isolation of human TWEAK receptor (TWEAKR)-expressing clones
XX from a COS cell human cDNA library. The TWEAK protein is a
XX member of the tumour necrosis factor (TNF) family and induces
XX angiogenesis. TWEAKR may therefore be used to screen for and develop
XX TWEAK agonists and antagonists for the modulation of angiogenesis, to be
XX used in the treatment and diagnosis of human disease. The disorders
XX mediated by angiogenesis include ocular disorders characterised by ocular
XX neovascularisation such as diabetic retinopathy, neovascular glaucoma,
XX retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
XX rubecosis, uveitis, macular degeneration and corneal graft
XX neovascularisation, and inflammatory diseases such as arthritis,
XX rheumatism and psoriasis. Other treatable diseases include malignant and
XX metastatic conditions such as sarcomas and carcinomas, benign tumours and
XX preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
XX scleroderma, vascular adhesions, atherosclerotic plaque
XX neovascularisation, telangiectasia, wound granulation, coronary
XX atherosclerosis, peripheral atherosclerosis and ischaemia.
XX
XX Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other:
XX
XX
XX alignment_scores:
XX Quality: 1066.00 Length: 207
XX Ratio: 5.150 Gaps: 0
XX Percent Similarity: 100.000 Percent Identity: 100.000
XX
XX alignment_block:
XX US-09-245-198A-4 x AAS03964 ..
XX
XX Align seg 1/1 to: AAS03964 from: 1 to: 898
XX
XX 78 SerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluCl 94
XX |||||||
XX 250 ACTTTGGGGAGCCGGGCGCATCGCTGCGCCCGAGGAGCTGGCCAGAGGA 299
XX |||||||
XX 94 uLeuValAlaGluGluValAspGlnAspProSerGluLeuAsnProGlnTrng 111
XX |||||||
XX 300 GCTGGTGCACAGAGGAGACGAGACCCGTCGGAATCGATCCCGACAGC 349
XX |||||||
XX 111 IuGluSerGlnAspProAlaProPheLeuAsnArgLeuValArgProArg 127
XX |||||||
XX 350 AAGAAAGCCAGATCTCGGCGCTTCTCGAAGACGACTGTTGGCGCTCGC 399
XX |||||||
XX 128 ArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIleAlaAl 144

```



```

400 AGAAGTGCACCTAAAGCCGCGAAACACGGGCTCGAAGACGATCGCAGC 449
144 AHIATYGLIValHisProArgProGlyGlnAspGlyValaGlnIaGly 161
450 CCATTATGAACTCATCCACGACCTCGGACAGCAGCAGCAGCAGCAGT 499
161 aAlaSpGlyThrValSerGlyTrpGlnGlnAlaArgIleAsnSerSer 177
500 TGGACGGGACAGTGAAGTGGCTGGGAGGAGGAGCAGATCAACAGCTCC 549
178 ProLeuArgTrpAsnArgGlnIleGlyGlnPheIleValThrArgAla 194
550 CCTCGGCTGCTCAACCCGACATCGGAGGAGTATTATAGCACCCGGGCT 599
194 YLeuTYrTYrLeuTYrCysGlnValHisPheAspGlnGlyLysAlaVal 211
600 GCTCTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 649
211 YrLeuLysLeuAspLeuValAspGlyValIleValAlaLeuArgCysLeu 227
650 ACCGGAAGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 699
228 GlnGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArg 244
700 GAGGAATTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCT 749
244 uCysGlnValSerGlyLeuAlaAlaLeuArgProGlySerSerLeuArg 261
750 CTGCGAGGTGCTGGGCTGGTGGCTGGGCTGGGCTGGGCTGGGCTGG 799
261 LeuArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThr 277
800 TCCGACACCTCCCTCGGCGCCATCTCAAGGCTGCCCTTCTCACTAC 849
278 PheGlyLeuPheGlnValHis 284
850 TTCGACTCTCTCCAGGTTTCAC 870

```

seq_name: /SDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:AAx23424

seq_documentation_block:

ID AAX23424 standard; DNA; 1030 BP.

XX AAX23424;

DT 18-JUN-1999 (first entry)

XX Human TNRL3 DNA.

KM Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KM developmental abnormality; gestational abnormality; prostate cancer;
 KM APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KM apoptosis; human; ss.

XX Homo sapiens.

OS
 FH key Location/Qualifiers
 FT CDS 1..627
 FT /tag= a
 FT /product= "TNRL3"

XX MO9911791-A2.

XX 11-MAR-1999.

XX 04-SEP-1998; 98MO-US18393.

XX 05-SEP-1997; 97US-0924634.

XX (UNITW) UNITW WASHINGTON.

```

PI Chaudhary PM;
XX
DR WPI: 1999-205191/17.
DR P-PSDB; AAM93590.
XX
PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
XX developmental or gestational abnormalities
XX
XX Example VII; Fig 13A; 156bp; English.
XX
CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other:

```

alignment_scores:

Quality: 1062.00

Ratio: 5.130

Percent Similarity: 99.519

Percent Identity: 99.519

alignment_block:

US-09-245-198a-4 x AAX23424 ..

Align seg 1/1 to: AAX23424 from: 1 to: 1030

```

77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnI 93
1 GTCAAGTTTGGGGAGCCGGGCAATCGCTGTCGCCGACGAGCCTGCCAGCA 50
93 uGluLeuValAlaGlnGluAspGlnAspProSerGluLeuAsnProGlnT 110
51 GGACCTGCTGGCAGAGGAGCAGCAGGACCTCGCAATGATATCCCA 100
110 hrgGlnGluSerGlnAspProAlaProPheLeuAsnArgLeuValArgPro 126
101 CAGAAAGAACCCAGAGATCTCGCCTTCTCGAACCCAGCTGTCGGCCT 150
127 ArgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIleAl 143
151 CGCAGAAGTGCACCTAAAGCCGGAACACGCGGCTCGAAGAGCGATCGC 200
143 aAlaHisTYrGlnValHisProArgProGlyGlnAspGlyValaGlnIa 160
201 AGCCATTATGAACTCATCCACGACCTGGGAGGAGGAGCAGCAGCAG 250
160 lYValAspGlyThrValSerGlyTrpGlnGlnAlaArgIleAsnSerSer 176
251 GTGTGAGCGGAGCAGTGAAGTGGCTGGGAGGAGGAGCAGATCAACAGCTCC 300
177 SerProLeuArgTrpAsnArgGlnIleGlyGlnPheIleValThrArgAl 193
301 AGCCCTGCTGCTACTACCGCAGATCGGAGGAGTATTATAGCACCCGGGC 350

```

193 acGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlnGlyLysAlaVal 210
|||||
351 TGGGCTCTACTACTGTACTGTACAGTGCACCTTGATGAGGGAAAGCTG 400
210 aTyrLeuLysLeuAspLeuValAspGlyValLeuAlaLeuArgCys 226
|||||
401 TCTACCTGAAGCTGACCTTGTGTGGATGTGTGTGCTGCTGCTGCTGCTG 450
227 LeuGlnGlnPheSerAlaThrAlaHisSerLeuGlyProGlnLeuArg 243
|||||
451 CTGGAGGAATTTCTCAGCCACTGCGGCACTTCCCTGCGGCCAGCTCCG 500
243 gLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeu 260
501 CCTGTGCCAGGTGTCTGCGCTGTGGCCCTGCGGCCAGGGTCTCCCTGCG 550
260 rGlnLeuArgThrLeuProThrAlaHisLeuLysAlaAlaProLeuThr 276
|||||
551 GGATCGCGACCTCCCTGCGGCCCACTCAAGGCTGCCCTTCTCTCACCC 600
277 TyrPheGlyLeuPheGlnValHis 284
|||||
601 TACTTCGGACTCTTCCAGGTTCAAC 624
seq_name: /SIDSL/gcdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV18599
seq_documentation_block:
ID AAV18599 standard; cDNA: 1168 BP.
XX
AC AAV18599:
XX
DT 21-JUL-1998 (first entry)
XX
DE Mus musculus tumour necrosis factor related ligand (TRELL) gene.
XX
KW TRELL; tumour necrosis factor related ligand; tnfr; treatment;
KW cancer; autoimmune disease; immune system; stimulation; suppression;
KW graft rejection; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT 2..679
FT /*tag= a
FT /note= "tumour necrosis factor related ligand"
XX
MO9805783-A1.
12-FEB-1998.
XX
PF 07-AUG-1997; 97WO-US13945.
XX
PR 18-MAR-1997; 97US-0040820.
PR 07-AUG-1996; 96US-0023541.
PR 18-OCT-1996; 96US-0028515.
XX
PA (BIOJ) BIOGEN INC.
PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.
PI Browning JL, Chicheportliche Y;
XX
DR WPI: 1998-145619/13.
DR P-PSDB: AAM47524.
XX
PT Tumour necrosis factor related ligand - useful for, e.g. treating
PT cancer, auto-immune disease and immune responses to tissue grafts
XX
PS Claim 2; Pages 45-46; 69pp: English.
XX
CC The sequence is that encoding mouse tumour necrosis factor related
CC ligand (TRELL). TRELL or active fragments can be included with a
CC carrier in pharmaceutical compositions to treat cancer, autoimmune

CC diseases or immune responses to tissue grafts, or to stimulate or
CC suppress the immune system. It is useful to screen for TRELL
CC receptors, by labelling with a detectable label and screening
CC compositions for binding. Agents interfering with TRELL-receptor
CC binding can also be screened for, can then be administered,
CC optionally with interferon- gamma, to induce cell death or
CC treat, suppress or alter immune responses (especially involving human
CC adenocarcinoma cells) involving a signal pathway between TRELL and its
CC receptor. The DNA sequence can be used in gene therapy for
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
CC autoimmune and inflammatory diseases or inherited genetic disorders,
CC by introducing into cells, and expressing, therapeutically effective
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
CC It may also be of use in the preparation of prepare probes for
CC screening natural/synthetic DNAs for TRELL-encoding sequences
CC and for antisense therapy.
XX
SO Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other:
alignment_scores:
Quality: 1020.00 Length: 224
Ratio: 4.744 Gaps: 0
Percent Similarity: 95.982 Percent Identity: 88.839
alignment_block:
US-09-245-198A-4 x AAV18599 ..
Align seg 1/1 to: AAV18599 from: 1 to: 1168
61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValVa 77
|||||
5 CTGAGCCTGGCCCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 54
77 IserLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGlu 94
|||||
55 CAGCCTGGGAGCTGGGCAACGCTGTCTCCAGGAGCTTCTTCAGAGG 104
94 IuLeuValAlaGlnGlnAspGlnAspProSerGlnLeuAsnProGlnThr 110
|||||
105 AGCTGACACAGACAGACCGCCGGAGCCCTCGAAGTGAATCCAGACA 154
111 GlnGlnSerGlnAspProAlaProPheLeuAsnArgLeuValArgProAr 127
|||||
155 GAGGAAGCAGAGATGTGTACTTCTTGAACAACACTAGTCGGGCTCG 204
127 gArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIleAla 144
|||||
205 AAGAACTGCTCTAAAGCGGGAAGCGCGCTCGCGAGCTATTGCG 254
144 IahIstYrGluValHisProArgProGlyGlnAspGlyValaGlnaIaGly 160
|||||
255 CCCATTATGAGGTTCATCTCGGCCAGACAGATGAGACACAGAGGT 304
161 ValAspGlyThrValSerGlyTyrGlnGlnAlaArgIleAsnSerSerse 177
|||||
305 GTGGATGGGACAGTAGTGTGGGAAAGACCAAAATCAACAGCTCCAG 354
177 rProLeuArgTyrAsnArgGlnIleGlyLubPheIleValThrArgAlaG 194
|||||
355 CCTGTGCGCTAGACCGCCAGATTGGGAATTTACAGCATCAGAGCTG 404
194 ILeuTyrTyrLeuTyrCysGlnValHisPheAspGlnGlyLysAlaVal 210
|||||
405 GGCTCTACTACTGTACTGTACAGTGCACCTTGATGAGGAAAGCTGTGC 454
211 TyrLeuLysLeuAspLeuValAspGlyValaLeuAlaLeuArgCysLe 227
|||||
455 TACCTGAAGCTGAGCTGTGTGGAAGCGGTGTGCGCTCGCTGCTGCT 504
227 uGlnGlnPheSerAlaThrAlaHisSerLeuGlyProGlnLeuArgL 244
|||||
505 GGAAGAAATTCAGCCACAGCAAGCAAGCTCTCTGCGGCCAGCTCGTT 554

244 eucysglnvalserglyleuleuvalalearproglyserleuarg 260
 |||||
 555 TGTGCCAGGTGTCTGGGCTGTGGCGGCGGAGGCTCTCCCTTCGG 604
 261 Tleargthleuprotpralaahisleuysalaalaproheleuthr 277
 |||||
 605 ATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCCCTTCTTAACCTA 654
 277 rphleglyleupheglnvalhis 284
 |||||
 655 CTTTGACCTCTTCACAGTTCAC 676
 seq_name: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:AAV23425

seq_documentation_block:
 ID AAX23425 Standard; DNA; 701 BP.

AC AAX23425;

18-JUN-1999 (first entry)

Mouse TNRL3 DNA.

Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 developmental abnormality; gestational abnormality; prostate cancer;
 APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 apoptosis; mouse; ss.

Mus sp.

| Key | Location/Qualifiers |
|--------|---------------------|
| FT CDS | 1..636 |
| FT | /tag= a |
| FT | /product= "TNRL3" |

W09911791-A2.

11-MAR-1999.

04-SEP-1998; 98WO-US18393.

05-SEP-1997; 97US-0924634.

(UNITW) UNIV WASHINGTON.

Chaudhary PM;

WPI: 1999-205191/17.

P-PSDB: AAM93591.

New Tumor Necrosis Factor family receptor polypeptides and ligands -
 useful for diagnosis and treatment of prostate cancer and
 developmental or gestational abnormalities

Example VII: Flg 13B: 156bp: English.

This invention describes isolated Tumor Necrosis Factor (TNF) family
 receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 their active fragments. APO4 is useful for diagnosing prostate cancer
 by determining levels of APO4 in an individual. Prostate cancer can also
 be treated using APO4 selective binding agents linked to a therapeutic
 moiety. APO4 polypeptides are also useful for identifying selective
 binding agents, useful in diagnosis/treatment of disease by binding of
 agents to the polypeptide/active fragment which is extracellular, or
 expressed on the cell surface. The binding is preferably performed in
 vivo. APO4 polypeptides/active fragments are also useful for screening
 for agonists and antagonists by binding and observing the change in APO4
 activity. Effective pharmacological agents useful in diagnosis or
 treatment of disease are also identified using APO4 polypeptides/active
 fragments and APO4 signal transducer molecules that specifically interact

CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.

Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

| alignment_scores: | Quality: | Ratio: | Percent Similarity: | Length: | Gaps: | Percent Identity: |
|-------------------|----------|--------|---------------------|---------|-------|-------------------|
| 968.00 | 4.768 | 96.209 | 89.100 | 211 | 0 | |

alignment_block:
 US-09-245-198a-4 x AAX23425 ..

Align seg 1/1 to: AAX23425 from: 1 to: 701

74 leuvalaivalserglyserargalaserleuarglnpr 90
 |||||
 1 CTGGTCGTGTACGCTGGGAGCTGGGCAAGCTGTCTGCCAGAGCC 50
 90 calaglglucluleuvalaialaglgluaspglaspserleuarg 107
 |||||
 51 TTTCAGAGAGAGCTGACACAGAGACCGCGGAGACCCCTCACTGA 100
 107 snprglnthrglucluleuaspglaspalaproheleuarg 123
 |||||
 101 ATCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 150
 124 valargproargseralaprolyglarygshrrargalarg 140
 |||||
 151 GTCCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 200
 140 galalialaialahistrgluvalahisproargproglylasp 157
 |||||
 201 AGCTATTCAGAGCCATATGAGGTTCATCTCCGCGCAGAGAGATGAG 250
 157 laglnalaglvalaspglythrvalserglytrpglnalarg 173
 |||||
 251 CACAGCAGAGGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 174 asnserserserproleuargtyrasnarglnleuyluphele 190
 |||||
 301 AACAGCTCAGAGCCCTGTGCTACGAGCCAGATTGGGCAATTACAG 350
 190 lthargalaglyleuargtyrtyrleuargtyrtyrleuarg 207
 |||||
 351 CATCAGAGGCTGGGCTCTACTACCTGCTACTGCTCAGGTGCTATGAG 400
 207 lylysalaivaltyrleuysleuaspgleuvalaspglyval 223
 |||||
 401 GAAAGGCTGTCTACTGAGCTGAGCTGCTGAGAGAGGTCTGCGCC 450
 224 leuargcysleuclulupheseralethralalaserleu 240
 |||||
 451 CTGGCTGCTGAGAGAGATTCTCAGCCACAGCAGAGAGAGAGAG 500
 240 oglnleuargleucysglnvalserglyleuvalalearpro 257
 |||||
 501 CACAGCTCGTTTGTGCGAGGTGTGGGCTGTGGGCTGGGCGAGGT 550
 257 erserleuarglileargthleuprotpralaahisleuysala 273
 |||||
 551 CTTCCTTCGAGATCCGACACCTCCCTGGGCTCATCTTAAGGCTG 600
 274 pheleuthrtyrphleglyleupheglnvalhis 284
 |||||
 601 TTCTTAACCTACTTGTGAGCTCTTCAAGTTCAC 633
 seq_name: /SIDSI/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:AAV24140

```

seq_documentation_block:
ID AAV24140 standard; DNA: 2856 BP.
XX
AC AAV24140:
XX
DT 28-SEP-1998 (first entry)
XX
DE Homo sapiens BARD1 exon 1.
XX
KW BARD1; ring protein; BRCA1; breast cancer; risk; diagnosis; exon; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH exon 2031..2188
FT /*tag= a
FT /number= 1
XX
W09812327-A2.
XX
26-MAR-1998.
XX
19-SEP-1997; 97WO-US16842.
XX
PR 04-APR-1997; 97US-0042985.
PR 20-SEP-1996; 96US-0025296.
PR 03-APR-1997; 97US-0042611.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Baer R, BOWCOCK AM;
XX
DR WPI: 1998-230317/20.
XX
PT DNA sequence encoding BARD1, B123, BE2, BE14, BE31 or BE445 - which
PT as breast cancer antigen, BRCA1, binding proteins are useful to
PT identify patient having or at risk of developing cancer
XX
PS Claim 25; Page 307-308; 348pp; English.
XX
CC The sequence is that of containing a BARD1 exon which can be used in the
CC preparation of the recombinant breast cancer antigen, BRCA1, binding
CC proteins BARD1, B123, BE2, BE14, BE31 or BE445, or a composition for the
CC detection of a BARD1, B123, BE2, BE14, BE31 or BE445 nucleic acid
CC sequence, specifically a wild type BARD1 composition for the detection
CC or purification of BRCA1, useful to identify a patient having, or at
CC risk of developing cancer. BARD1 can be used in the preparation of an
CC anti-BARD1 antibody, and in the detection and purification of a BRCA1
CC protein. BARD1, B123, BE2, BE14, BE31 or BE445 can be used in the
CC identification of a binding protein agonist or antagonist that alters
CC the binding of BARD1, B123, BE2, BE14, BE31 or BE445 to BRCA1 or the
CC biological activity of the BRCA1-BARD1, B123, BE2, BE14, BE31 or BE445
CC complex. The antibodies can be used to detect BARD1, B123, BE2, BE14,
CC BE31 or BE445, a specific anti-BARD1 antibody can be used to identify
CC a patient having or at risk of developing cancer.
XX
SQ Sequence 2856 BP; 616 A; 778 C; 786 G; 674 T; 2 other;
XX
alignment_scores:
Quality: 127.00 Length: 325
Ratio: 0.836 Gaps: 15
Percent Similarity: 46.769 Percent Identity: 25.846
alignment_block:
US-09-245-198A-4 x AAV24140 ..
Align seg 1/1 to: AAV24140 from: 1 to: 2856
14 ProLeuPtoArgSerLeuGlySerItrGAspGlyAlaIVal..... 27
|||||
|||||
|||||
1788 CCCACATCGGCACGCGCTTGTCTGTGCGGCCCCCGCCCTGTGCGCCCTGCGCA 1837

```

| | | | |
|------|---|--|------|
| 28 | | ArgIAlaGlnPro..... | 32 |
| 1838 | GTCCCTATTTTGGAGACATGTGGGCGCGCGTCCCGCCCTCCCGCCG | 1887 | |
| 33 | ..ProAlaProMetAlaAlaArgArgSerGlnArgArgArgIArgArg | 48 | |
| 1888 | ACCCGCGCCCTGTGGCGCCCGCCCTCCACAGCGGGAAGAGCTTGCC | 1937 | |
| 49 | GlyGluProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLe | 65 | |
| 1938 | GGTTTCGAGTCCCTGCGCTGACGCTTCCCT..... | 1967 | |
| 65 | uAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValaValSerLeuGlySera | 82 | |
| 1968 | | GTGGTTTCCCGAGGCGTCC | 1986 |
| 82 | rgAlaSerLeuSerAlaGlnGluProAlaGlnGlnGluLeuValAlaGlu | 98 | |
| 1987 | TTCCTTCCTCCGCTCTCGAGACCTTTTATCTCGAAGCGCGGACATGCG | 2036 | |
| 99 | GluAspGlnAspProSerGluLeuLeuAsnProGlnThrGlnGluSerGlnAs | 115 | |
| 2037 | GATTAATCGGCACACCGAGAACCGGACCGCAGAGATCCGCTCCGGGAACA | 2086 | |
| 115 | pPro.....AlaProPheLeu.AsnArgLeuValArgProArgArgSer | 129 | |
| 2087 | GCCTCGTTCGCGCGCCCGGCATGGAACCGAGATGTCGCGCTCTGGGCC | 2136 | |
| 130 | AlaProGlyGlyArgIArgIThrArgIArgIArgAlaIleAlaIleHisI | 146 | |
| 2137 | ACAGTCGCGCCCGCGCTCGACCGCTGAGAAAGCTGCTGCGTCT.... | 2181 | |
| 146 | rgLuValHisProArgProGlyGlnAspGlyAlaGlnAlaCly..... | 160 | |
| 2182 | | CGCGTTTGGTAAAGACGAGAGCTCTTGGGGGTGGCGTC | 2218 |
| 161 | ..ValAspGlyThrValSerGlyIArgIArgIAlaArgIleAsnSerSer | 176 | |
| 2219 | GAGGCGACGCGGTGCGACAGATTCTGTGGGCGGACAAATCTTTCAAAAT | 2268 | |
| 177 | SerProLeuArgIArgIAsnArgIle..... | 185 | |
| 2269 | TCCGTTTCCCTCTTCCGCTCCGCGCTCGACAGCGGGTGGGCGTGCAGTA | 2318 | |
| 186 |GlyGluPheIleValIThrArgAlaClyLeuIArgIArgILeuT | 199 | |
| 2319 | GCACCTGCGCGGGGAGTATAGTATTAACAATCTTCGCTCTCATCAGCTT | 2368 | |
| 199 | ysCysGlnValHisPheAspGluGlyLysAlaValTyr.....Leu | 212 | |
| 2369 | ATTTTTGG..... | GGACATTACGGCGTCCCTT | 2397 |
| 213 | LysLeuAspLeuLeuValaAspGlyValaLeuAlaLeuArgCysLeuGluG | 229 | |
| 2398 | GCCTCAATCCAACTGAAACGGTAGTATTGTAACCTTCAGCGGCTGAGA | 2447 | |
| 229 | upPheSerAlaThrAlaAlaSerSerLeu.GlyProGlnLeuArgLeuCys | 245 | |
| 2448 | ACCGTGTGAGGTGTGGCTTAAAAAATGTCTTCCCGGCTCGCGCTCCAG | 2497 | |
| 246 | GlnValSer.....GlyLeuLeuAlaLeuArgProGlu | 256 | |
| 2498 | AGTTTCAATCGGCGCTGGGCTGGGCGTGAAGCTTTCGATTTTAAACCGG | 2547 | |
| 256 | ysSerSerLeuArgIleArgIThrLeuPro.....TrrAla. | 267 | |
| 2548 | CCGTGATTAACCC...CGCTGCTTCCGCGGAGCTGTGCGAATTGGCGTGC | 2594 | |
| 268 |HisLeuLysAlaAlaProPheLeuThr | 276 | |
| 2595 | GGCGGCGCGCCCGGAGACCTCTTAATTAAGAAGCGGTCCGACACTCTTAAG | 2644 | |

277 Tyrophegilyleupheginal 283
|||||
2645 TAAACTGGCCTTTTGACATT 2655

seq_name: /STD1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAA49205

seq_documentation_block:
ID AAA49205 standard; cDNA; 1660 BP.

AC AAA49205;

DT 12-DEC-2000 (first entry)

DE Corn putative lecithin:cholesterol acyltransferase gene #5.

KM Corn: lecithin:cholesterol acyltransferase; phytoosterol;
XX phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock; ss.

OS Zea mays.

Key Location/Qualifiers

CD5 77.1396

FT /tag= a
FT /product= "lecithin:cholesterol acyltransferase"
FT /EC_number= "2.3.1.43"

PN W0200032791-A2.

PD 08-JUN-2000.

PF 02-DEC-1999; 99WO-US28586.

PR 03-DEC-1998; 98US-0110782.

PA (DUPO) DU PONT DE NEMOURS & CO E. I.

PI Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JJ;

DR WPI: 2000-412337/35.

DR P-PSDB; AAB01210.

XX Polynucleotide encoding plant lecithin:cholesterol acyltransferase
XX enzyme useful for producing transgenic plants and for producing
XX antibodies specific to which is useful for screening cDNA expression
XX libraries -

XX Claim 2: Page 43-44; 49pp: English.

XX The present sequence is a putative coding sequence for a corn
XX lecithin:cholesterol acyltransferase (also known as
XX phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
XX associated with high-density lipoproteins and apolipoprotein A1 and -D.
XX The gene and protein can be used to produce transgenic plants which have
XX increased lipid metabolism and membrane fluidity, and therefore increased
XX resistance to heat and/or cold shock, to alter the content of phytoosterol
XX or lecithin in grains and to identify potential herbicides.

XX Sequence 1660 BP; 361 A; 475 C; 499 G; 325 T; 0 other;

alignment_scores:

Quality: 126.50 Length: 319

Ratio: 0.866 Gaps: 17

Percent Similarity: 45.768 Percent Identity: 25.392

alignment_block:

US-09-245-198A-4 x AAA49205 ..

Align seg 1/1 to: AAA49205 from: 1 to: 1660

11 ArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAlaVal 27
:::|||||
394 CAACGAGTACCGGAACCTCGCCGG.....CGTCGAGAGCGGAGTGCACA 437

27 1.....ArgGlnAlaGlnProProAlaPro.....M 36
438 ACTTCGGCTCCACAGAGACCTTCAGCCACAGAAACCCCTCAATCAGAC 487
36 etAlaAlaArgArgSerGlnArgArgGlyArgArgGlyProGly 52
:::|||||
488 TGTGTCTCGGAAGCTGAGAGCCGACCTGAGACATGATGATGACCGAGA 537
53 ThrAlaLeuLeuVal.....ProLeuAlaLeuGlyLeuGlyLeuAla 67
538 CGAGAGACCATGTTCCGAGCCCTCAGACTTCGCTACCGCCGCCGT 587
67 u.....AlaCysLeuGlyLeuLeuLeuAlaValAla 78
588 CCCCCGGCCAGACGTCGAGGTGATCTCCGCTTCAAGAGACTGATG 637
78 erLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGlu 94
638 GAGCTGCTCA..... 648
95 LeuValAlaGlnGluAspGlnAspProSerGlnLeuAsnProGlnThr 111
649GGCCGCGAGCGAGAGACCG.....GA 671
111 uGlySerGlnAspProAlaProPheLeuAsnArg.....Leu 124
672 AGAAGCCGCTATCTCTGCGCACAGCTTCGGCGCATGTCGCCCTGAG 721
124 AlArgProArgArgSerAlaProGlySerGlyArgGlyThrArgAlaArg 140
|||||
722 TTCTCCGGAACACTCCGCC.....GGCGTGGCGCGGAGCAGACATCA 765
141 AlAlaAlaAlaHisTyrGluValHisProArgProGlyGln..... 154
766 GCG.....CTTCGTCTGTGTCGCCGCGCAGCC 791
155AspGlyAlaGln..... 158
792 TCCCCGGCGGTTCTCGAGCCGCTGCGCACACTTCGCCGTCGCGAGCAG 841
159AlaGlyValAspGlyThrValSer 166
842 ATCTTACAGTCCAGGAGCAGCCGCTGCGCACGCGAGCAGCATGTGAG 891
167 .GlyTyrGluGluAlaArgIleAsnSerSerProLeuArgTyrAsnA 183
892 GAGCTCGAGAGCCCATCGGAACCTCCCTCCGCGCGCTGTCGGCC 941
183 rGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyr 199
942 GCTTCAGGCGCCCTGCTGTCACAGAGCGGAACTAC..... 982
200 CysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAsp 216
983TCCGCGTCCGCGCAGCAGCATGAGGCG 1008
216 uLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAla 233
1009 CTTCCTCGCCGCGCTCGGCTCCGCGAGCGCGCGAGCCCTTCAGAGAC 1058
233 hrAlaAlaSerSerLeuGly.....ProGlnLeuArgLysCys 245
1059 GGGCCGTCGCCAAGATGGCAGCTTCGCGCGCGCGGATGTCATACG 1108
246 GlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIle 262
1109 TACATCAGCGGGGTC.....GGCAACAGAGAGCGCGTGGC 1143
262 gThrLeuProThrAlaHis.....LeuLysAlaAlaProPheLeuThrTyr 278
1144 GCTGTGTCTTGGGGCGAAGACTTCGACGCGCGCCGAGAGTGGCGGT 1193

278 hegly 279

:::|
1194 ACGGG 1198

seq_name: /SISL/gcdata/geneseq/geneseqn-embL/NA1999.DAT:AAV69289

seq_documentation_block:

ID: AAV69289 standard; cDNA: 2832 BP.

AAV69289;

01-MAR-1999 (first entry)

Murine JNK-Interacting protein 1 (JIP-1) cDNA.

JIP-1: JNK-Interacting protein; c-Jun NH2-terminal kinase;
signal transduction; inhibitor; mouse; neurodegenerative disease;
Parkinson's disease; Alzheimer's disease; blood clot; stroke;
malignancy; cancer; leukaemia; autoimmune disease; inflammation;
apoptosis; therapy; diagnosis; ds.

Mus sp.

Location/Qualifiers

180..2162

/*tag= a

W09849188-A1.

05-NOV-1998.

28-APR-1998; 98WO-US08513.

28-APR-1997; 97US-0819177.

(UYMA-) UNIV MASSACHUSETTS.

Davis RJ, Dickens M;

WPI: 1999-024042/02.

P-PSDB: AAW81525.

c-Jun NH2-terminal kinase (JNK)-interacting protein 1 - used to
treat neurodegenerative disease, blood clot, leukaemia, autoimmune
disease, and inflammation

Claim 11: Page 66-70; 95pp; English.

This cDNA sequence codes for murine JNK-interacting protein 1
(JIP-1, see AAW81525), a novel cytoplasmic anchor protein that
specifically binds to and inhibits the biological effects of JNK
(c-Jun NH2-terminal kinase), including the initiation of apoptosis
and oncogenic transformation. To isolate the clone, a yeast
two-hybrid method was used to screen a mouse embryo cDNA library to
identify proteins that interact with JNK. 7 overlapping partial
clones were isolated, and the full-length cDNA was subsequently
obtained by screening a mouse brain lambda ZAPII cDNA library. The
invention provides JIP-1 nucleic acids and polypeptides, expression
vectors and host cells. The JIP-1 polypeptide and nucleic acids
(including antisense and ribozyme molecules) can be used in the
manufacture of a medicament for treating a pathological condition
associated with abnormal expression or activity of JNK, such as a
neurodegenerative disease (selected from Parkinson's disease and
Alzheimer's disease), a blood clot, stroke, malignancy, leukaemia,
an autoimmune disease or inflammation (all claimed).

Sequence 2832 BP; 569 A; 888 C; 840 G; 535 T; 0 other;

alignment_scores:

Quality: 124.50

Ratio: 1.064

Percent Similarity: 44.487

Length: 263

Gaps: 12

Percent Identity: 25.475

alignment_block:

US-09-245-198a-4 x AAV69289

Align seg 1/1 to: AAV69289 from: 1 to: 2832

```
1 MetSerLeuLeuAspPheGluIleSerAlaArgArgLeuProLeuProAr 17
:::|||||:
20 ATAACTTGATATCGCTGCGAGACCGG.....CCGCCCCAGC 60
17 gSerLeuGlySerArgAspGlyAlaValArgGlnAla..... 30
:::
61 TCAGTCCGAACCCGCGGCGGCGGCTCTCCACACGCGCTCCGCTGC 110
31 .....GlnProAlaProMetaAlaAlaArg 40
|||||:
111 TCCCGCGCGCGCGCGCGCGCGCTCCGCGCGCTCCGCGCGGATG 160
41 SerGln.....ArgArgArgGlyArgArgGlyGluProGlyThrAl 54
:::
161 GCCAGGCGCTGCACCCCGGAAATGGCGAGAGAGAGAGCGGCTGGCGGG 210
54 aleuLeuValProLeuAla.....LeuGlyLeu..... 63
:::
211 GCGCGCGCTCCCGCGCGCGCTCCCGCTCCCGCTCCCGCTCCCGCT 260
63 ..... 63
261 TCCGCTCCCAATTTCAGGCTCACCACATGATCAGCTGAGAGATTGCA 310
64 .....GlyLeuAlaLeuAlaC 69
311 GGATGAAGACCTTTCGAGATCAGTACAGATGTGATCAGCTGAGT 360
69 ys.....LeuGlyLeu.....LeuLeuAlaVal 76
:::
361 GCAAGACACCTGCTCTCGCGCGCGCGCGCGCGCGCTGCTGCTGG 410
77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGln 93
|||
411 GGTACAGCGGCGCGCGCGGAGCGGCTGCAGGCGGAGATCTGCAGAT 460
93 uGluLeuVal.....AlaGluGlnAspGln 102
:::
461 GGACCTGCTCGACGCGCGCGCTGACACTCCGCGCGCGCGAGAGAG 510
102 sp.....ProSerCLeuLeuAspPro 108
:::
511 AGGAGAGAGAGAGAGAGCTGCTGCCCAAGACAGAGAGAGAGAGCTCC 560
109 GlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeuValAr 125
:::
561 AAGCGGAGATCCACACGAGATCCGGGCT.....CG 592
125 gProAlaArgSerAlaPro.....LysGlyArgGlyThrArgAla 139
:::
593 CAGCGAGGCGCGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 642
139 rArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 155
643 AGAGGCTTACACGCTCAACCTTTTCCCGAGAGTCCGCGGCTCAGAGC 692
156 GlyAlaGlnAlaGlyValAspGlyThrValSerGlyThrGluAlaIleAr 172
:::
693 ACGCTGAATTAATTAATCTTTAGCAAAAGACACAGTGGCAGAGACCGT 742
172 gIleAsnSerSerSerProLeuArgTyrAsnArgGln 184
:::
743 GTCTGATCATCTCCCTCTGAAGACAGAGAGAGAGAG 779
```

seq_name: /SISL/gcdata/geneseq/geneseqn-embL/NA2002.DAT:AAU25519

seq_documentation_block:


```
seq_name: /STD01/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAA09686
seq_documentation_block:
ID AAA09686 standard; DNA: 3957 BP.
XX
AC AAA09686;
XX
D7 31-JAN-2001 (first entry)
XX
DE HSV-2 immediate early protein ICP4 DNA sequence.
XX
KW Herpes-simplex-virus type 2; HSV-2; Infected cell protein 4; ICP4;
RV vaccine; Infection; ds..
XX
OS Herpes simplex virus type 2.
XX W09516779-A1.
PD 22-JUN-1995.
XX PF 13-DEC-1994: 94MO-EP04138.
XX PR 14-DEC-1993: 93GB-0025496.
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Pala P., Cheysen DR., Slaout MM., Koutsoukos MC.;
XN WPJ.: 2001-024142/03.
DR P-PSDB: AAB26874.
PT Immediate early herpes-simplex-virus type 2 (HSV-2) ICP4 protein is
PT used in vaccines for therapeutically or prophylactially treating HSV
PT infections -
XX Claim 5; Page 16; 28pp; English.
XX PS CC This invention relates to an immediate early herpes-simplex-virus type 2
XX CC (HSV-2) infected cell protein 4 (ICP4) recognised by human cytotoxic T
CC cells. HSV-2 ICP4 protein is recognized by cytotocic T-lymphocyte (CTL)
CC cells in humans and is used in vaccines for therapeutically or
CC prophylactially treating HSV infections. Pharmaceutical compositions of
CC HSV-2 ICP4 protein may be used to treat patients suffering from HSV
CC infections, to prevent or decrease recurrent herpes disease, frequency,
CC severity and duration of episodes. The present sequence represents HSV-2
DNA encoding ICP4.
XX Sequence 3957 BP: 368 A; 1656 G; 1568 C; 365 T; 0 other;
XX
Alignment_scores:
Quality: 124.00 Length: 171
Ratio: 1.610 Gaps: 8
Percent Similarity: 45.029 Percent Identity: 30.409
alignment_block:
US-09-245-198A-4 x AAA09686 ..
Align seg 1/1 to: AAA09686 from: 1 to: 3957
10 AlaAtgArGLeuProleuProArgserLenglysSerAgAspGlYAl 26
||||||| ||| ||| :::: | || | |||
GCCCGCCCCGGCCC CGC GCCGC GGCGGCCGCGGG...CGCGCCGCGCCG 2053
26 aValARglInAGlnPrOALePRoMetAlalalaIarg..... 39
|::::::::::| :||| ||| :::||| |||
CGCACGCCAACGCCGCCCGCCCTGCCGCGGTGCTGCGGACTGCGGGTTC 2103
2054 ..... ArgSer 41
```

```

2104 GTCCGCGACGCGCTGCTGCTGATCTGCGCTGCGCGGAGCACTGCGCGCTGGC 2153 |||||
42 GlnArgArgArgGlyArgGly.....GluProGlyTh 53 |||||
2154 CGCGCGGACGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2203 |||||
53 rAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysL 70 |||||
2204 CCGG.....GGCCCTGGGCGCGC..... 2222 |||||
70 euGlyLeuLeuLeuAlaValValSerLeuGlySerArgAlaSerLeuSer 86 |||
2223 .....GCT 2225 |||
87 AlaGlnGluProAlaGlnGluLeuValAlaGlnGluAspGlnAspR 103 |||||
2226 GCGCGGAGCGCGCGCGCTGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGAGC 2275 |||||
103 oSerGlyLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheL 120 |||||
2276 TGCTCTTCGACCA.....GACCTCGCGCGCGCGCTG 2307 |||||
120 euAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLysThr 136 |||||
2308 CTGG.....CGACCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2348 |||||
137 ArgAlaArgArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 153 |||||
2349 CGCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2395 |||||
153 yGlnAspGlyAla 157 |||||
2396 CCAGGCGCGCGCGC 2408 |||||

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.AAD2519
seq_documentation_block:
ID AAD2519 standard; DNA; 154746 BP.
AC
AA AAD2519:
XX
DT 26-MAR-2002 (first entry)
DE
XX
XX Human herpesvirus 2 complete DNA genome.
XX
XX Human herpesvirus 2; cytostatic; cancer; immunosuppressive; virucide;
XX antiherpetic; fungicidal; protozoicidal; antirheumatic; antiinflammatory;
XX antiallergic; rheumatoid arthritis; neuroprotective; multiple sclerosis
XX immune response; vasotropic; vaccine; gene therapy; autoimmune disease;
XX vasculitis; ds.
XX
XX Human herpesvirus 2.
XX
XX WO200176643-A1.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001MO-US11372.
XX
XX 07-APR-2000; 2000US-19560P.
XX
XX (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
XX Orson FM, Kinsey BM, Bhogal BS;
XX
XX WPI; 2002-066308/09.
XX
XX
XX Composition for oral delivery of vaccines, comprises expressed vector
XX containing antigenic genomic sequence, bound to aggregated
XX protein-polycationic polymer conjugate or suspension
XX
XX Disclosure; Page 90-132; 145bp; English.
XX

```


XX The invention relates to a composition comprising an expression vector
CC bound to an aggregated protein-polycationic polymer conjugate or
CC suspension. The expression vector contains a promoter polynucleotide
CC sequence operatively linked to a polynucleotide sequence encoding an
CC antigen which is a fragment of a gene or genome associated with an
CC infectious disease, cancer and autoimmune disease such as rheumatoid
CC arthritis, vasculitis, and multiple sclerosis, pathogenic genomes
CC consisting of bacterium, fungus, protozoa and virus such as human
CC immunodeficiency virus (HIV), herpes simplex virus (HSV), hepatitis C
CC virus (HCV), influenza and respiratory syncytial virus (RSV), and
CC optionally comprising a nucleotide sequence encoding a cytokine (or a
CC cytokine expression vector), is useful for inducing an immune response
CC (systemic and/or mucosal) in an organism. The cytokine expression vector
CC contains a sequence for granulocyte macrophage-colony stimulating factor
CC (GM-CSF) or interleukin-12 (IL-12). The polynucleotide sequences encoding
CC the antigen and the cytokine are under transcriptional control of same or
CC different promoter polynucleotide sequences. The expression vector, as a
CC DNA vaccine is useful for treating a condition in an organism. The
CC present sequence is human herpesvirus 2 complete DNA genome related
CC to the invention.

SQ Sequence 154746 BP; 23003 A; 54218 C; 54701 G; 22824 T; 0 other;

alignment_scores:

Quality: 124.00 Length: 171
Ratio: 1.610 Gaps: 8
Percent Similarity: 45.029 Percent Identity: 30.409

alignment_block:

US-09-245-198A-4 x AAD25519 ..

Align seg 1/1 to: AAD25519 from: 1 to: 154746

```
10 AlaArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAla 26
   |||||  |||  |||  ::  |||  |||  |||
151716 GCCCGCGCGCGCGCGCGCGCGCGCGCGCGG...CGCGCGCGCGCGCGC 151762

26 aValArgGlnAlaGlnProProAlaProMetAlaAlaArg..... 39
   |::|::|::|  |||||  ::|::|::|::|
151763 CGCAGCGCGCGCGCGCGCGCGCTGCGCGCTGCGGAGCTCGCGGTTTC 151812

40 .....ArgSer 41
   |||::|
151813 GTGGCGAGCGCGCTGCTGATGCGCTGCGCGGAGACCTGCGCGTGGC 151862

42 GlnArgArgArgGlyArgArgGly.....GluProGlyTh 53
   ::|::|::|::|  |||||  |||||
151863 CGCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCTGAGCTGATGTCG 151912

53 rAlaLeuLeuValProLeuAlaLeuGlyLeuAlaLeuAlaCysL 70
   ||  ||  ||  ::
151913 CCGG.....GGCCCTGGGCGCGCGG..... 151931

70 eugGlyLeuLeuLeuAlaValValSerLeuGlySerArgAlaSerLeuSer 86
   ::|
151932 .....GCT 151934

87 AlaGlnGluProAlaGlnGluGluLeuValAlaGlnGluAspGlnAspPr 103
   |||  |||||  |||||  ::::|  ::|
151935 GCCCGGAGCGCGCGCTGCTGAGCTCGCGCGCGCGCGCGCGGAGAC 151984

103 oSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheL 120
   |::|  ::|::|
151985 TGCTCTTCACAGACCA.....GAGCGCTGCGCGCGCTG 152016

120 euAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLysThr 136
   ::|  |||||  ::|  ::|::|::|::|
152017 CTGGC.....CGACACCGCTGCGCGCGCGCGCGCGCTGCGCGCGCC 152057

137 ArgAlaArgArgAlaAlaAlaAlaHisTyrGluValHisProArgProG1 153
   |||  |||  ::|::|::|  ::|  ::|::|::|::|
```

```
152058 CGCCTCCGCGCGCGCGGAGCGCGCAAGCGCA...GAGCCCGCGCGCGG 152104
153 yGlnAspGlyAla 157
   ||||  ::|::|
152105 CCAGGCGCGCGCGC 152117
```

THIS PAGE BLANK (USPTO)

OM of: US-09-245-198a-4 to: GenEmbl:* out-format: pfs
Date: Jun 22, 2002 2:30 PM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=frim+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09245198/runat.21062002.184006.9945/app_query.fasta.1.632  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPOP=4.500  
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000  
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DEEXT=7.000 -START=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEADSIZE=500  
-MINLEN=0 -MAXLEN=200000000 -USER=US09245198 -ECGN1_1.10597  
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30  
-NO_XLPEXY -WAIT -THREADS=1
```

Search information block:

Query: US-09-245-198a-4
Query length: 284
Database: GenEmbl:*
Database sequences: 1797656
Database length: 187333701
Search time (sec): 2600.680000

score_list:

| Sequence | Strid | Orig | Zscore | Escore | len | Documentation |
|-----------------|-------|---------|---------|---------|--------|---------------------------------|
| gb_pat:AX201324 | + | 1366.00 | 1243.78 | 6.5e-61 | 1353 | AX201324 Sequence 3 from Patent |
| gb_pr:AF055872 | + | 1366.00 | 1243.69 | 6.5e-61 | 1368 | AF055872 Homo sapiens Apoc/DR3 |
| gb_pr:BC019047 | + | 1331.50 | 1211.08 | 4.3e-59 | 1651 | BC019047 Homo sapiens, Simlial |
| gb_pr:AF030099 | + | 1298.00 | 1152.65 | 1.6e-57 | 1306 | AF030099 Homo sapiens, TWEAK m |
| gb_pat:AR140407 | + | 1265.00 | 1153.27 | 7.1e-56 | 1236 | AR140407 Sequence 1 from Patent |
| gb_pat:AX180714 | + | 1066.00 | 976.06 | 5.3e-46 | 898 | AX180714 Sequence 1 from Patent |
| gb_ro:AF030100 | + | 1020.00 | 932.48 | 1.4e-43 | 1168 | AF030100 Mus musculus TWEAK m |
| gb_hlg:AC016876 | - | 589.50 | 504.24 | 1.0e-13 | 19038 | AC016876 Homo sapiens c10ne |
| gb_hlg:AC069459 | - | 555.50 | 473.03 | 5.5e-18 | 203083 | AC069459 Mus musculus c10ne |
| gb_ro:AL603707 | - | 555.50 | 471.93 | 6.3e-18 | 234182 | AL603707 Mouse DNA sequence |
| gb_hlg:AC098923 | - | 385.50 | 320.50 | 1.7e-09 | 179030 | AC098923 Mouse musculus c10ne |
| gb_pl:AF002482 | + | 145.50 | 103.41 | 2.1e+03 | 187835 | AF002482 Oryza sativa genome |
| gb_da:SC1C2 | - | 140.00 | 103.03 | 912.82 | 42210 | AL133124 Streptomyces coelic |
| gb_da:AL646059 | - | 138.50 | 97.00 | 4.9e+03 | 190050 | AL646059 Ralstonia solanace |
| gb_hlg:AC105470 | + | 136.50 | 93.59 | 7.5e+03 | 233912 | AC105470 Rattus norvegicus c |
| gb_da:AF002222 | + | 135.00 | 120.15 | 249.27 | 6397 | AF002222 Pseudomonas aerugin |
| gb_da:AE004651 | - | 135.00 | 116.14 | 416.92 | 10728 | AE004651 Pseudomonas aerugin |
| gb_da:AE004740 | - | 134.00 | 115.06 | 478.40 | 10965 | AE004740 Pseudomonas aerugin |
| gb_da:SCC61A | - | 130.00 | 101.35 | 2.8e+03 | 40352 | AL335695 Streptomyces coelic |
| gb_da:SCDMSSTS | + | 129.00 | 115.56 | 449.22 | 5751 | X89010 S. glaucus genes str |
| gb_da:SCC46 | + | 129.00 | 107.53 | 1.3e+03 | 16185 | AL139164 Streptomyces coelic |
| gb_da:SCG46 | - | 129.00 | 104.02 | 2.0e+03 | 25459 | AL006985 Streptomyces glauc |
| gb_da:SCG46 | - | 129.00 | 104.02 | 2.5e+03 | 32941 | AL133469 Streptomyces coelic |
| gb_da:SCM10 | + | 129.00 | 99.78 | 3.4e+03 | 43961 | AL133422 Streptomyces coelic |
| gb_pr:CMS01DV6 | - | 129.00 | 90.01 | 1.2e+04 | 137833 | AL133455 Human chromosome 14 |
| gb_pr:AC016708 | - | 127.00 | 86.45 | 1.9e+04 | 194304 | AC016708 Homo sapiens BAC c1 |
| gb_da:AL646063 | + | 126.50 | 85.50 | 2.1e+04 | 207050 | AL646063 Ralstonia solanace |
| gb_pr:HSBARD11 | + | 125.50 | 85.97 | 118.10 | 999 | AF038034 Homo sapiens BRC1-ass |
| gb_da:AVIN1FC | + | 125.50 | 99.90 | 3.3e+03 | 28795 | M20568 A. vinelandii mfor nlf |
| gb_hlg:AP003822 | - | 125.50 | 91.31 | 1.0e+04 | 87143 | AF003822 Oryza sativa chromo |
| gb_hlg:AP004272 | + | 125.50 | 86.91 | 1.8e+04 | 153595 | AP004272 Oryza sativa chromo |
| gb_da:SCG42 | - | 125.00 | 98.98 | 3.8e+03 | 30590 | AL031371 Streptomyces coelic |
| gb_da:SCG610 | - | 124.50 | 97.11 | 4.8e+03 | 36734 | AL049497 Streptomyces coelic |
| gb_hlg:AP003876 | - | 124.50 | 89.46 | 1.3e+04 | 98504 | AP003876 Oryza sativa chromo |
| gb_hlg:AP003911 | - | 124.50 | 88.81 | 1.4e+04 | 107132 | AP003911 Oryza sativa chromo |
| gb_v1:HSV2HG52 | - | 124.50 | 85.95 | 2.0e+04 | 154746 | 286099 Herpes simplex virus |
| gb_hlg:AP003634 | + | 124.50 | 84.93 | 2.3e+04 | 176580 | AP003634 Oryza sativa chromo |
| gb_pat:AA5258 | + | 124.00 | 113.94 | 552.59 | 3957 | AA5258 Sequence 2 from Patent |
| gb_pr:HS745C22 | - | 124.00 | 104.17 | 1.9e+03 | 13949 | AL031596 Human DNA sequence |
| gb_v1:HSV2HG52 | + | 124.00 | 85.50 | 2.1e+04 | 154746 | 286099 Herpes simplex virus |

seq_name: gb_pat:AX201324

| seq_name | LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS | TITLE | JOURNAL | FEATURES | BASE COUNT | ORIGIN |
|-----------------|----------|-----------------------------------|-----------|-------------|----------|--------|--------------|---------------------|---|--|----------------------|----------|------------|--------|
| gb_pat:AX201324 | AX201324 | Sequence 3 from Patent WO0153486. | AX201324 | GI:15391154 | | human. | Homo sapiens | 1 (bases 1 to 1353) | Askenazi, A.J., Goddard, A., Godowski, P.J., Gurney, A.L., Hillan, K.J., Marsters, S.A., Pan, J., Pitti, R.M., Roy, M.A., Smith, V., Stone, D.M., Watanabe, C.K. and Wood, W.I. | Compositions and methods for the treatment of tumour | Genentech, Inc. (US) | 257 a | 443 c | |

seq_documentation block:

| LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS | TITLE | JOURNAL | FEATURES | BASE COUNT | ORIGIN |
|----------|-----------------------------------|-----------|-------------|----------|--------|--------------|---------------------|---|--|----------------------|----------|------------|--------|
| AX201324 | Sequence 3 from Patent WO0153486. | AX201324 | GI:15391154 | | human. | Homo sapiens | 1 (bases 1 to 1353) | Askenazi, A.J., Goddard, A., Godowski, P.J., Gurney, A.L., Hillan, K.J., Marsters, S.A., Pan, J., Pitti, R.M., Roy, M.A., Smith, V., Stone, D.M., Watanabe, C.K. and Wood, W.I. | Compositions and methods for the treatment of tumour | Genentech, Inc. (US) | 257 a | 443 c | |

alignment_scores:

Quality: 1366.00 Length: 268
Ratio: 5.097 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment block:

US-09-245-198a-4 x AX201324 ..

Align seg 1/1 to: AX201324 from: 1 to: 1353

| | | |
|-----|---|-----|
| 17 | ArgSerLeuGlySerArgSpGlyGlyAlaValArgGlnAlaGlnProPr | 33 |
| 1 | CGATCCCTGGGTCCCGGATGGGGGGCGGTGAGGACAGACAGCCCC | 50 |
| 33 | oAlaProMetaAlaAlaArgArgSerGlnArgArgGlyArgGlyG | 50 |
| 51 | CGCCCGCAGCGCCCGCCGCGAGCCAGAGCGGAGCGCGCGCGG | 100 |
| 50 | LnProGlyThrAlaLeuValProLeuAlaLeuGlyLeuAla | 66 |
| 101 | ACCAGGCGACCCCGCTGCGTCCGCTGCGCTGCGCTGCGCTGCG | 150 |
| 67 | LeuAlaCysLeuGlyLeuLeuLeuAlaValAlaSerLeuGlySerArg | 83 |
| 151 | CTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 200 |
| 83 | aSerLeuSerAlaGlnGluProAlaGlnGluLeuValAlaGlnGlu | 100 |
| 201 | ATCGCTGCTGCGCCAGAGAGCGCCAGAGAGAGAGAGAGAGAG | 250 |
| 100 | spGlnAsProSerGluLeuAsnProGlnThrGlnGlnSerGlnAsPro | 116 |
| 251 | ACCAGAGACCGCTGCGAATCCCGACAGAGAGAGAGAGAGATCCT | 300 |
| 117 | AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGly | 133 |
| 301 | GGCGCTTCTTACCGACTAGTTGCGCTGCGAGAGAGAGAGAGAG | 350 |
| 133 | YArgLysThrArgAlaArgArgAlaIleAlaIleAlaIleAlaIle | 150 |

```
351 CCGGAAAACAGGGCTCGAAGAGGATCGACCCCATTTATGAAGTTCATC 400
150 roatpTroglyInaSpGlyAlaGlnaIaGlyValaIaSpGlyThrValSer 166
|||||
401 CAGGACTGGACAGACGAGCGAGCGAGCTGTGGACGGGACAGTACT 450
167 GLYTTPGluGlnaIaArgIleasnSerSerProleuArgTyrAsnAr 183
|||||
451 GCGTGGAGAGCCAGATCAACAGCTCCAGCCCTCGCTGAGTACACCG 500
183 ggInIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrC 200
501 CCAAGATCGGGAGATTATAGTACCCCGGGCTCTCTACTACTACTACT 550
200 ysgInValHisPheaspGluGlyLysAlaValTyrLeuLysLeuAspLeu 216
|||||
551 GTGAGGTGCACTTTGATGAGGAGGAGCTGTACTCTGAAGCTGACTTG 600
217 LeuValaSpGlyValLeuAlaLeuAlaArgCysLeuGluGluPheSerAlaTh 233
|||||
601 CTGGTGGATGTTGTGCTGCGCTGCTGCTGAGGAATTCACAGCCAC 650
233 rAlaIaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyL 250
|||||
651 TGGGGCGAGTTCCTCGGGCCCAAGCTCCCTCTGCGAGGTGTCTGGC 700
250 euleuAlaIaLeuArgProGlySerSerLeuArgIleArgThrLeuProTyr 266
|||||
701 TGTGGGCGCCCGGCGCCAGGGTCCCTCGGGAGTCCGACCCCTCCCTGG 750
267 AlaHisLeuLysAlaIaIaProPheLeuThrTyrPheGlyLeuPheGlnVal 283
|||||
751 GCCCATCTCAAGGCTGCCCTCTCTCTCCTACTCTGAGACTCTTCACAGT 800
283 LHis 284
|||||
801 TCAC 804
```

seq_name: gb_pr:AF055872

seq_documentation_block:

LOCUS AF055872 1368 bp mRNA linear PRI 04-MAY-1998

DEFINITION Homo sapiens ApO3/DR3 ligand (APO3L) mRNA, complete cds.

ACCESSION AF055872

VERSION AF055872.1 GI:3108230

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 1368)

Marsters,S.A., Sheridan,J.P., Pittl,R.M., Brush,J., Goddard,A. and

Ashkenazi,A.

IDENTIFICATION OF A LIGAND FOR THE DEATH-DOMAIN-CONTAINING RECEPTOR

apO3

JOURNAL Curr. Biol. 8 (9), 525-528 (1998)

MEDLINE 98228355

REFERENCE 2 (bases 1 to 1368)

Marsters,S.A., Sheridan,J.P., Pittl,R.M., Brush,J., Goddard,A. and

Ashkenazi,A.

Direct Submision

Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way,

South San Francisco, CA 94080, USA

FEATURES

source

1..1368

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="17"

/map="17p13"

1..1368

/gene="APO3L"

58..807

/gene="APO3L"

```
/function="binds to the death-domain-containing receptor
ApO3/DR3"
/Note="TNF homolog"
/codon_start=1
/product="ApO3/DR3 ligand"
/protein_id="AAC39724.1"
/db_xref="GI:3108231"
/translation="MAARSORRRGRGRRGPGTALLVPLALGLALGLALLAVSL
GSRSLASNOEPAOEELVAEEDODSELPOTEEESODPAFLNRLVRRPRSPKGRKTR
ARRAIAHVEVHPRGODCAQAGVDTGTSWGEARINSSPLRYNROIGERTVRAGL
YVLQVHFDESKAVYLLKDLVNGVALRLRLEFSRATMASSLPQLRLCOVSLIAL
RPGSLRLRLTLPMAHLKAAPFLYFGLFQVH"
BASE COUNT 272 a 443 c 389 g 264 t
ORIGIN
```

```
alignment_scores:
Quality: 1366.00 Length: 268
Ratio: 5.097 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

```
alignment_block:
US-09-245-198a-4 x AF055872
Align seg 1/1 to: AF055872 from: 1 to: 1368
```

```
17 ArgSerLeuGlySerArgAspGlyGlyAlaValArgGlnaIaGlnProPr 33
|||||
1 CGATCCCTCGGATCCCGGATGGGGGGCGGTGAGCAGGACACAGCCCC 50
33 oAlaProMetaIaIaIaArgArgSerGlnaArgArgGlyArgArgGly 50
|||||
51 CGCCCCATGGCGCCCGCTCGAGACAGAGCGAGGGGGCGCGGGGGGG 100
50 LuProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAla 66
|||||
101 AGCCGGGACCGCCCTGCTGCTCCCGCTCGCTGGGCTGGGCTGGCG 150
67 LeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAl 83
|||||
151 CTGGGCTTCTCGGCTCCCTGCTGCGCGGTGCTGAGTTTGGGGAGCCGGC 200
83 aSerLeuSerAlaGlnaIaProAlaGlnaIaGlnaIaLeuAlaGlnaIa 100
|||||
201 ATGCGCTCCGCGCCAGAGGCTGCCAGAGAGCTGTGTCAGAGAGAGG 250
100 spGlnaSerProSerGlnaLeuAsnProGlnThrGluGluSerGlnaAsp 116
|||||
251 ACCAGGACCGGTGGAACTGAATCCCGACAGAGAAAGCCAGAGATCT 300
117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysG 133
|||||
301 GCGGCTTCTCGTGAACGACAGTACGTCGCGCTCGCAGAAAGTCCACTTA 350
133 yArgLysThrArgAlaArgArgAlaIleAlaIaHisTyrGluValHisP 150
|||||
351 CCGGAAAACACGGGCTCGAAGAGCGATCCGACGCCATTATGAAGTTCATC 400
150 roatpTroglyInaSpGlyAlaGlnaIaGlyValaIaSpGlyThrValSer 166
|||||
401 CAGGACTGGAGAGGAGCGAGCGAGCGAGCTGTGGACGGGACAGTAGT 450
167 GLYTTPGluGlnaIaArgIleasnSerSerProleuArgTyrAsnAr 183
|||||
451 GCGTGGAGAGCCAGATCAACAGCTCCAGCCCTCGCTGAGTACACCG 500
183 ggInIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrC 200
501 CCAAGATCGGGAGATTATAGTACCCCGGGCTCTCTACTACTACTACT 550
200 ysgInValHisPheaspGluGlyLysAlaValTyrLeuLysLeuAspLeu 216
|||||
551 GTGAGGTGCACTTTGATGAGGAGGAGCTGTCTACTGAGAGCTGAGACTTG 600
```

217 leuValaspGIyValleuAlaleuArqCysleuGlugluPheseralath 233
|||||
601 CTGGTGAAGGTGTCTGTGGCCCTGGCCCTGCTGAGAGATTCACACCAC 650
233 rAlaAlaSerSerleuGIyProGlnleuArqleuCySglnValSerGIyL 250
|||||
651 TCGGGGAGATCTCCCTCGGGCCCGACGCTCCGCTTCGACAGTGTCTGGGC 700
250 leuAlaAlaLeuArqProGIySerSerleuArqGlnleuArqProtrp 266
|||||
701 TGTGGCCCTGGGGCCAGGCTCTCCCTCGGATCCGACCCCTCCCTGG 750
267 AlAHsleuLysAlaAlaIaProPheleuThrTyPheGIyLeuPheGlnVa 283
|||||
751 GCCCATCTCAAGGCTGCCCTTCTCTACCTACTTGGACTTCCAGGT 800
283 lHs 284
|||||
801 TCAC 804
seq_name: gb_pr:BC019047
seq_documentation_block:
LOCUS BC019047 1651 bp mRNA linear PRI 11-DEC-2001
DEFINITION Homo sapiens, Similar to tumor necrosis factor (ligand)
superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
complete cds.
ACCESSION BC019047.1 GI:17512138
VERSION BC019047.1
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1651)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapsb-remail.nih.gov
Tissue Procurement: Louis Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@hgtl.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantirpop, S., Thomas, P.J.,
Tongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, D.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 30 Row: P Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4507596.
Location/Qualifiers
1. 1651
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:20669 IMAGE:4766071"
/tissue_type="Primary B-Cells from Tonsils"
/clone_id="NIH_MGC_48"

/lab_host="DH10B-R"
/note="Vector: pOTB7"
106.510
CDS
/codon_start=1
/product="Similar to tumor necrosis factor (ligand)
superfamily, member 12"
/protein_id="AAH19047.1"
/db_xref="GI:17512139"
/translation="MAARRSRRRGRGREGTALLVPLALGLALACGLLAAVSL
GSRASLSAOPAEELVAEDQDPSLNPOTESQDPAPPLNLRVPRRSAPGKTR
ARRAIAAHYEHHPGGDGAADADGGYTCARP"
BASE COUNT 344 a 517 c 481 g 309 t
ORIGIN
alignment_scores:
Quality: 1331.50 Length: 356
Ratio: 4.84 Gaps: 2
Percent Similarity: 77.247 Percent Identity: 77.247
alignment_block:
US-09-245-198a-4 x BC019047 ..
Align seg 1/1 to: BC019047 from: 1 to: 1651
9 SerAlaArqArqleuProleuProArqSerleuGIySerArqaspGIyGL 25
|||||
25 TCGGCCCGCGGCTCCCTCCCTCCCGATCCCTGGGTCCCGGATGGGG 74
25 yAlaValArqGlnAlaGlnProProAlaProMetaAlaAlaArqArqSerg 42
|||||
75 GCGGTGAGAGGACGACACGCCCGCCCGCATGGCGCGCGCGAGCG 124
42 lArqArqArqGIyArqArqGIyGIyProGIyThAlaLeuLeuValPro 58
|||||
125 AAGAGCGAGGAGGCGCCCGGGGGAGACCGGACCGCTCGTGTGCCG 174
59 leuAlaLeuGIyLeuGIyLeuAlaLeuAlaCysleuGIyLeuLeuLeuAl 75
|||||
175 CTGGCGCTGGCTGGGCTGGGCTGGGCTGGCTGGCTGGCTGGCTGGG 224
75 aValAlaSerleuGIySerArqAlaSerleuSerAlaGlnGIyProAlaG 92
|||||
225 CGGTGTCATTTGGGGAGCGGCGCATCGCTGCCCGCCAGAGCGTGGC 274
92 lnglGIyLeuValAlaGIyGIyAspGlnAspProSergIyLeuAsnPro 108
|||||
275 AGGAGAGCTGGTGGCGAGAGAGACAGACCGCTCGGAATGATCCC 324
109 GlnThrGIyGlnSerGlnAspProAlaProPheleuAsnArqleuValAr 125
|||||
325 CAGACGAGAGAAAGCCAGATCCCTGGCTTCTGACCGATGATTTGG 374
125 gProArqArqSerAlaProIyGIyArqIyThrArqAlaArqArqAlaI 142
|||||
375 GCCTGCCAAGAAGTGCACCTAAAGCGCGGAAACACGGCTCGAAGGCCA 424
142 leAlaAlaHisTyGIyValAlaHisProArqProGIyGlnAspGIyAlaGln 158
|||||
425 TCGACACCCATTATGAAATTCAATCCAGACCTGACAGACGAGAGCGCAG 474
159 ..Ala..... 159
475 GCAGATGAGAGTTTACACAACCTTGTGAGGCCATGAGATCTAAGTGGTG 524
159 159
525 GAGCCAAAGATTGAACCAAGTAGAATGTGCTGTACTGGACATGTGTC 574
159 159
575 TGCATGAAGAGGAGGTGCTGTGTGACAGGGGTGAGGGGTCCATGCAAGGG 624

159 159
625 CCACATCCAAAAAGGAGAGAGAAAGTTCCAGAAAGAGAGACACA 674
160 G1yValAs 162
675 TCCTCCACCATTTACCAGAGGTCGAAGAGACAAAGACACCAGGTGGGA 724
162 pG1yThValSerG1yTTPG1yG1yValA1aArg1yLeAsnSerSerProL 179
725 CGGAGACAGTGAAGTGGCTGGAGGAGGAGGAGATCAACAGCTCCAGCCCTC 774
179 euArg1yTyrAsnArg1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1y 195
775 TGCCCTACACACCCAGATCGGGAGATTATAGTCACCCGGGCTGGGCTC 824
196 TTTTyrLeuTyrCysG1yValHisPheAspG1yG1yValA1aTyrLeu 212
825 TACTACTGTAAGTCAAGTGCACCTTGTATGAGGAGGAGGCTGCTACTC 874
212 uLysLeuAspLeuValAspG1yValA1aLeuValA1aLeuArgCysLeuG1y 229
875 GAACCTGGACTGCTGGTGGATGGTGGCTGGCCCTGGGCTGGAGG 924
229 1uPheSerAlaThrAlaAlaSerSerLeuG1yProG1yLeuArgCys 245
925 AATTCTCAGCCACCTCGCGGAGTTCCCTCGGGCCCGACGCTCGGCTGTC 974
246 G1yValSerG1yLeuLeuAlaLeuArgProG1ySerSerLeuArg1yLeu 262
975 CAGGTGCTGGGCTGTGGCCCTCGGGCAGGCTCCTCGCGGAGTCG 1024
262 gThLeuProTPrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheG 279
1025 CACCTCCCGGGGCGCATCTCAAGGCGCCCTTCCTCCTCAGCTTTCG 1074
279 1yLeuPheG1yValHis 284
1075 GACTCTTCAGGTTCCAC 1091
seq_name: gb_pr:AF030099
seq_documentation_block: 1306 bp mRNA linear PRI 20-DEC-1997
LOCUS AF030099
DEFINITION Homo sapiens TWEAK mRNA, complete cds.
ACCESSION AF030099
VERSION AF030099.1 GI:2707218
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1306)
Chicheportiche, Y., Bourdon, P.R., Xu, H., Hsu, Y.M., Scott, H.,
Hession, C., Garcia, I. and Browning, J.L.
TWEAK, a new secreted ligand in the tumor necrosis factor family
that weakly induces apoptosis
J. Biol. Chem. 272 (51), 32401-32410 (1997)
98070415
2 (bases 1 to 1306)
Bourdon, P., Hession, C., Tizard, R. and Browning, J.
Direct Submission
Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
Cambridge, MA 02142, USA
FEATURES
source
1..1306
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/feature_type="cns11"
18..767
/note="Ligand in the TNF family; secreted protein; start

codon not verified experimentally"
/codon_start=1
/product="TWEAK"
/protein_id="AAC51923.1"
/db_xref="GI:2707218"
/translation="MAARSSQRRRGRRRPEPTALLVPLALGLALACTGLLAVSL
GSRASLSQEPQAEELVAEEDDPSELNPQTEESDPAFLNRLVRRPRSPKGRKTR
ARRAIAHYEVHPRPGDQAOAGVDGVSGWEMEARINSSPLRVNRQIGETIVRAGL
YLYCOVHEDEKAVYLRDLVDVGVLRCLIEERSATVAASLGPQLRCOVSGLLAL
RPGSLRIRTRTPMAHLKAPFLTYRGLPOVH"
BASE COUNT 247 a 434 c 368 g 257 t
ORIGIN
alignment_scores:
Quality: 1298.00 Length: 254
Ratio: 5.110 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-245-198a-4 x AF030099 ..
Align seg 1/1 to: AF030099 from: 1 to: 1306
31 G1nProProAlaProMetAlaAlaArgArgSerG1nArgArgG1yArg 47
3 CAGCCCCCCCCCATGGCCGCGCCGCGAGCAGAGCGGAGGCGGCG 52
47 gArgG1yG1yProG1yThrAlaLeuValProLeuAlaLeuG1yLeuG 64
53 CCGGGGGAGCGGGGACCGCGCTGCTGCTGCGCTGCGCTGCGCTGCG 102
64 1yLeuAlaLeuAlaCysLeuG1yLeuLeuLeuAlaValA1aSerLeuG1y 80
103 GCTGGGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 152
81 SerArgAlaSerLeuSerArgAlaG1yG1yProAlaG1yG1yLeuVal 97
153 AGCGGGCATCGCTGCTGCGCCAGAGACCTGCGCAGAGAGAGCTGTCG 202
97 ag1yG1yAspG1yAspProSerG1yLeuAsnProG1yThrG1yG1ySerG 114
203 AGAGAGAGACAGAGACCGCTGGAACTGAATCCCAAGACAGAAAGACC 252
114 1nAspProAlaProPheLeuAsnArgLeuValArgProAlaArgSerAla 130
253 AGATCTCGCGCTTCTCTGTAACGCACTAGTTCGCTCGCCTCGCAAGTCA 302
131 ProLysG1yArg1yThrArgAlaArgAlaLeuAlaHisTyrG1 147
303 CCTAAAGCGCGAAACACGGGCTCGAAAGCGATCGCACCCCATTTATGA 352
147 uValHisProArgProG1yG1yAspG1yAlaG1yValA1aAsp1yT 164
353 AGTTCATCCAGACTGTGACAGCGAGCGCAGCGAGTGTGACGGGA 402
164 hrValSerG1yTTPG1yG1yValA1aArg1yLeAsnSerSerProLeuArg 180
403 CAGTGAAGTGGCTGGAGAGAACCGCAATCAACAGCTCCAGCCCTCTCGC 452
181 TyrAsnArgG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1y 197
453 TACAACCGCCAGATCGGGAGATTATAGTCAACCGGCTGGGCTCTACTA 502
197 rLeuTyrCysG1yValHisPheAspG1yG1yValA1aTyrLeuLys 214
503 CCTTACTGTCAGGTGCATTTGATGAGGAGGAGGCTGTACTCTGAAGC 552
214 euAspLeuLeuValAspG1yValLeuAlaLeuArgCysLeuG1yG1yPhe 230
553 TGGACTTGCTGGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 602
231 SerAlaThrAlaSerSerLeuG1yProG1yLeuArgLeuG1yG1yVal 247

```
|||||
603 TCAGCCACTGCGGCGAGTTCCTCGGGCCCGACCTCGGCTTGCCAGGT 652
247 1SerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThrL 264
653 GTCCTGGCGTGTGGCCCTCGGCGAGGATCTCCGCGGATCCGACGCC 702
264 euProThrAlaHisLeuLysAlaAlaProPheLeuThrThrPheGlyLeu 280
703 TCCCTCGGCGCCCTCAAGGCTGCCCTTCTCAGCTTACTTGACACTC 752
281 pheGlnValHis 284
753 TTCAGAGTTCAC 764

seq_name: gb_pat:ARI40407
seq_documentation_block:
LOCUS ARI40407 1236 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6207642.
ACCESSION ARI40407
VERSION ARI40407.1 GI:14482903
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Wiley/S.R.
TITLE Member of the TNF family useful for treatment and diagnosis of
JOURNAL Patent: US 6207642-A 1 27-MAR-2001;
FEATURES Location/Qualifiers
source 1. 1236
BASE COUNT 225 a 416 c 358 g 237 t
ORIGIN

alignment_scores:
Quality: 1265.00 Length: 249
Ratio: 5.080 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.598

alignment_block:
US-09-245-198a-4 x ARI40407 ..
Align seg 1/1 to: ARI40407 from: 1 to: 1236
36 MetAlaAlaArgArgSerGlnArgArgGlyArgArgGlyArgGlyArgProG 52
1 ATGGCGCGCGCGTGGGAGCCAGAAAGCGGAGGGCGCGGCGGAGCGCGG 50
52 yThrAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAla 69
51 CACCGCGCTGTGTGCTCCGCTGCGGCTGGGCTGGGCTGGGCTGGGCT 100
69 yLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAlaSerLeu 85
101 GCGTGGCGCTCTGCTGGCGCTGTGTGCTAGTTGGGAGCGGGGATGCTG 150
86 SerAlaGlnGluProAlaGlnGluGluValAlaGlnGluAspGlnAs 102
151 TCCGCGCCAGAGAGCTGCCAGAGAGAGCTGTGGCGAGAGAGAGACGAGA 200
102 pProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaPro 119
201 CCCGTCGGAACCTGAATCCCGACAGAAAGAACCGAGATCTCGCGCTT 250
119 heLeuAsnArgLeuValArgProArgSerAlaProGlyArgGlyArg 135
251 TCCTGAACCGAGCTAGTTCGCGCTCGAAGAGTGCACCTAAAGCGCGGAA 300
136 ThrArgAlaArgArgAlaIleAlaIleHisTyrGluValHisProArgPr 152
```

```
|||||
301 ACAGCGGCTCGAAGAGCGATTCGAGGCCATTATGAGTTCAATCAGACAC 350
152 ocGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrG 169
351 TGGACAGAGACGAGCGCAGAGGTGTGAGCGGAGACGACGAGTGGCTGG 400
169 lngluAlaArgIleAsnSerSerProLeuArgTyrAsnArgGlnIle 185
401 AGGAAGCCAGAAATCAACAGCTCCAGCCCTCTGGCTTACAAACGCCAGATC 450
186 GlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnA 202
451 GGGAGTTTATATGATCAACCGGCGTGGCTCTACTACTGCTAGCTAGGT 500
202 HisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValA 219
501 GCACCTTGATGAGGGAGAGCTGTCTAAGCTGAGCTTGCTGTGG 550
219 spGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235
551 ATGCTGTGCTGCGCTGCGCTGCGTGGAGAAATCTCAGCCACTGCGGCG 600
236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAl 252
601 AGTTCCTCGGGCGCCAGCTCCGCTCTGCGAGGTCTGTGGCTGTGGC 650
252 AleuArgProGlySerSerLeuArgIleArgThrLeuProThrAlaHisL 269
651 CCGCGCGCCAGAGGCTCTCCCTCGAGATCCGACCCCTCGGCGCCATC 700
269 euLysAlaAlaProPheLeuThrThrPheGlyLeuPheGlnValHis 284
701 TCAAGGCTGCGCCCTCTCCTCACTTCTGAGACTTTCAGAGTTCAC 747

seq_name: gb_pat:AX180714
seq_documentation_block:
LOCUS AX180714 898 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1 from Patent W00145730.
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 898)
AUTHORS Wiley/S.R.
TITLE Tweak receptor
JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;
FEATURES Location/Qualifiers
source 1. 898
CDS
1. 898
/organism="synthetic construct"
/cdb_xref="taxon:32630"
52..873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=1
/protein_id="CAC50004.1"
/db_xref="GI:15132571"
/translation="MATGSRPTSLIARGLICTPLDQSGSATSPDRKQIDKITEITS
KITYHENEIRARIKILIGERTSRSLGSRASAOEPADBEIYAEEDQDPELNPOTERS
ODPAFLNRLVRPRRSAPKGRKTRARAIANHEVHRPPQDGAQAGVDGTVSGWEA
RINSSPLRVNRQIGERTIVRAGLYLVYCOVHPDEGRAVYLKLDLVDGVIALRCLDE
FSATPAASLSLPOLRLCOVSGILARLPSSSLRIRTLIPAHAKAFLTYFGLFQVH"
```

```
alignment_scores:
Quality: 1066.00 Length: 207
Ratio: 5.150 Gaps: 0
```

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-245-198a-4 x AX180714 ..

Align seg 1/1 to: AX180714 from: 1 to: 898

```
78 SerLeuGlySerArgAlaSerLeuSerAlaGlnProAlaGlnGlnGln 94
|||||
250 AGTTGGGAGCCGGGCGATCGCTCGCCGAGAGGCGTGGCCGAGAGGA 299
|||||
94 ULLeuValAlaGlnGlnAspGlnAspProSerGlnLeuAsnProGlnThr 111
|||||
300 GCTGCTGGCAGAGGAGGAGCAGGACCGTGGAACTGAATCCCGAGACAG 349
|||||
111 LUGLUSerGlnAspProAlaProPheLeuAsnArgLeuValArgProArg 127
|||||
350 AAGAAACCCAGAGATCGTGGCTTCTCGAACCGACTAGTTGGCTCGC 399
|||||
128 ArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIleAlaAl 144
|||||
400 AGAAGTCACCTAAAGCCGGAACACGGGCTCGAAGAGGATCGCAGC 449
|||||
144 aHISlyrGlyValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 161
|||||
450 CCATTAAAGAGTTTCATCCAGCAGCTGGAGCAGGAGCGAGCAGCAGTG 499
|||||
161 aLAspGlyThrValSerGlyTyrPglGlnAlaAlaArgIleAsnSerSer 177
|||||
500 TGGACGGGAGCAGTGTGCTGGGAGGAGCAGACATCAACAGCTCCAGC 549
|||||
178 ProLeuArgTyrAsnArgGlnIleGlyGlnPheIleValThrArgAla 194
|||||
550 CCTCTGGCTACACACCGCAGATCGGGAGATTATACACCCGGCGTG 599
|||||
194 yLeuTyrTyrLeuTyrCysGlnValHisPheAspGlyGlyValAlaVal 211
|||||
600 GCTTACTACTGTGATCTGAGGTGACTTTGATGAGGGGAGGCTGTCT 649
|||||
211 yrLeuLysLeuAspLeuValAspGlyValLeuAlaLeuArgCysLeu 227
|||||
650 ACCTGAAAGCTGGACTGTGCTGGATGTGTGCTGGCTGGCTGGCTG 699
|||||
228 GlnGlnPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArg 244
|||||
700 GAGGAATTCTCAGCCACTGGCGCAGTTCCCTCGGGCCCCAGCTCCGCT 749
|||||
244 uCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArg 261
|||||
750 CTGGCAGGTGTGTGGCTGTGGCCCTGGCCGAGGCTCCCTCGCGGA 799
|||||
261 lEarGlyrLeuProTyrAlaHisLeuLysAlaAlaProPheLeuThrTyr 277
|||||
800 TCCGCACCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCCTCAGCTAC 849
|||||
278 PheGlyLeuPheGlnValHis 284
|||||
850 TTCGACTCTTCAGGTTTCC 870
|||||
```

seq_name: gb_ro:AF030100

seq_documentation_block:

LOCUS AF030100 1168 bp mRNA linear ROD 20-DEC-1997

DEFINITION Mus musculus TWEAK mRNA, partial cds.

ACCESSION AF030100

VERSION AF030100.1

KEYWORDS

SOURCE

ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1168)
Chicheportliche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,

Hession,C., Garcia,I. and Browning,J.L.
TWEAK, a new secreted ligand in the tumor necrosis factor family
that weakly induces apoptosis
J. Biol. Chem. 272 (51), 32401-32410 (1997)

REFERENCE 2 (bases 1 to 1168)

AUTHORS Chicheportliche,Y., Bixler,S., Tizard,R. and Browning,J.

TITLE Direct Submission

JOURNAL Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
Cambridge, MA 02142, USA

FEATURES Location/Qualifiers

source

1..1168

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_type="peritoneal macrophages"

<1..679

/note="ligand in the TNF family; secreted protein"

/codon_start=2

/product="TWEAK"

/protein_id="AAC53517.1"

/translation="VLSGLALACILLLVYSLGSMTLSAOPSEBELFAEDRREP

PELNQTEESDDVPELFLQVLRPRRSAPKGRARRAIAAHYEVHPGGDAGV

DGVSGMEETKINSSPLRYDRQIEFTYIRAGLYLYLCVHPDEGRKAVYIAKLDELIV

GVLAIRCLIEFSATRAASSPGQLRLQVSGLLPLRPSLSLIRTLPMANHLAAPELTV

FLFLQVH"

BASE COUNT

242 a 360 c 298 g 268 t

ORIGIN

alignment_scores:

Quality: 1020.00 Length: 224

Ratio: 4.744 Gaps: 0

Percent Similarity: 95.982 Percent Identity: 88.839

alignment_block:

US-09-245-198a-4 x AF030100 ..

Align seg 1/1 to: AF030100 from: 1 to: 1168

```
61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValVa 77
|||||
5 CTGAGCCTGGGCGCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 54
|||||
77 lSerLeuGlySerArgAlaSerLeuSerAlaGlnGlnProAlaGlnGln 94
|||||
55 CAGCTGGGAGCGTGGGCAAGCGTGTGCTGCCAGGACCTTTCAGAGG 104
|||||
94 lLeuValAlaGlnGlnAspGlnAspProSerGlnLeuAsnProGlnThr 110
|||||
105 AGCTGACAGCAGAGGACCGCGGAGCCCTGTAACGTAATCCCGACACA 154
|||||
111 GlnGlnSerGlnAspProAlaProPheLeuAsnArgLeuValArgProAr 127
|||||
155 GAGGAAGCCAGGATGTGTACCTTCTTGGAACAACCTAGCCGGCTCG 204
|||||
127 gArgSerAlaProLysGlyArgLysThrArgAlaArgArgAlaIleAla 144
|||||
205 AAGAAAGTCTCTAAAGCCGGAAGCGCGGCGCTCGCGCAGATTCAG 254
|||||
144 lArHISlyrGlnValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160
|||||
255 CCCATTATGAGGTTCACTCGCGCAGAGCAGATGAGCACAAGCAGGT 304
|||||
161 ValAspGlyThrValSerGlyTyrPglGlnGlnAlaArgIleAsnSerSer 177
|||||
305 GTGGATGGAGCAGTGAAGTGGCTGGGAGAGACCAAAATCAACAGCTCAG 354
|||||
177 rProLeuArgTyrAsnArgGlnIleGlyGlnPheIleValThrArgAla 194
|||||
355 CCTCTCGCTACGACCGCCAGATGGGGAATTTACAGTCAATCAGGCTG 404
|||||
194 yLeuTyrTyrLeuTyrCysGlnValHisPheAspGlnGlyLysAlaVal 210
|||||
```



```

|||||
405 GCCTACTACTGACTGTGACGTGACACTTGTATGAGGAAAGCTGTC 454
211 TTTTLeuysLeuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLe 227
455 TACCTGAAGCTGGACTGCTGTGTGAAGCTGTGCGCCCTCGCTGCTCCT 504
227 uGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgL 244
505 GGAAACAATTCACGACACAGCAAGCTCTCTGCGCCCGACGCTCCGTT 554
244 euGyGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArg 260
555 TGTGCAGAGTGTCTGGCGCTGTGCGCGACGAGGCTTCCTCCGTCGG 604
261 TLeaGThrLeuProThrAlaHisLeuysAlaAlaProPheLeuThrTy 277
605 ATCCGACCCCTCCCTCGGCTCATCTTAAGGCTGCCCTTCCTAACCTA 654
277 rPheGlyLeuPheGlnValHis 284
655 CTTTGACTCTTTCAGTTTCAAC 676

seq_name: gb_hlg:AC016876

seq_documentation_block:
LOCUS AC016876 190358 bp DNA linear HTG 03-SEP-2001
DEFINITION Homo sapiens clone RP11-186B7, WORKING DRAFT SEQUENCE, 16 unordered
pieces.
ACCESSION AC016876 GI:15421989
VERSION AC016876.4
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 190358)
REFERENCE Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome, clone RP11-186B7
unpublished
2 (bases 1 to 190358)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,
Perrera,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heathford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karais,A., Klein,J.,
Lehoczky,J., Lien,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,Y., Riley,R., Roy,A., Santos,R., Sevely,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teste,S., Tittrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submmission
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02144, USA
On Sep 3, 2001 this sequence version replaced g1:13431059.
All repeats were identified using RepeatMasker:
Shult, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3849
Center clone name: 186_B.7
----- Summary Statistics
Sequencing vector: M13; M7815; 3% of reads
Sequencing vector: Plasmid; n/a; 97% of reads

```

```

Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 181928 bases at least Q40
Consensus quality: 185529 bases at least Q30
Consensus quality: 187335 bases at least Q20
Insert size: 168000; agarose-fp
Insert size: 188858; sum-of-coverage
Quality coverage: 8.1 in Q20 bases; agarose-fp
Quality coverage: 7.2 in Q20 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
20703: contig of 20703 bp in length
20704 20803: gap of 100 bp
20804 21935: contig of 1132 bp in length
21936 22035: gap of 100 bp
22036 24202: contig of 2167 bp in length
24203 24302: gap of 100 bp
24303 28199: contig of 3897 bp in length
28200 28299: gap of 100 bp
28300 35021: contig of 6722 bp in length
35022 35121: gap of 100 bp
35122 40745: contig of 5624 bp in length
40746 40845: gap of 100 bp
40846 47136: contig of 6291 bp in length
47137 47236: gap of 100 bp
47237 58980: contig of 11744 bp in length
58981 59080: gap of 100 bp
59081 68164: contig of 9084 bp in length
68165 68264: gap of 100 bp
68265 94625: contig of 26361 bp in length
94626 94725: gap of 100 bp
94726 102564: contig of 7839 bp in length
102565 102664: gap of 100 bp
102665 116764: contig of 14100 bp in length
116765 116864: gap of 100 bp
116865 136408: contig of 19544 bp in length
136409 136508: gap of 100 bp
136509 156380: contig of 19872 bp in length
156381 156480: gap of 100 bp
156481 180631: contig of 24151 bp in length
180632 180731: gap of 100 bp
180732 190358: contig of 9627 bp in length.
Location/Qualifiers
1. 190358
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-186B7"
/clone_11b="RPC1-11 Human Male BAC"
1. 20703
/note="assembly-fragment
clone_end:sp6
vector_side:left"
20804. 21935
/note="assembly-fragment"
22036. 24202
/note="assembly-fragment"
24303. 28199
/note="assembly-fragment"
28300. 35021
/note="assembly-fragment"
35122. 40745
/note="assembly-fragment"
40846. 47136
/note="assembly-fragment"
47237. 58980
/note="assembly-fragment"
59081. 68164

```



```

66609 TCAGGCTGATGGAGACAGTGCCTGGAGACACCAAAATCAACAG 66560
175 rrsertprrleuArgrtYrAsnargrInlleglygnllepelleValThra 192
66559 CTCACACCCCTGCTGCTACGACCGCAGATTGGGAAATTACAGTATCA 66510
192 rgalaglyleuTyrrYrleuTyrrCys.gln..... 201
66509 GGGCTGGGGCTACTACTGACTGTACAGTAAGCCCTGGCTGGCTTCAG 66460
201 ..... 201
66459 GGGCAGAGCAAAAGCTAAGGGAGAGAGAGACTGGCAAGAAATGGTGGCA 66410
201 ..... 201
66409 GAGGAGACTTCGGACTCATGAGAGAGATGCTGGCTTTGATGAGGG 66360
202 .....ValHisPheas 205
66359 AGGCAGAGCTGGATTGTGCTTCCTCTCTCTTCAGGTGCACCTTGA 66310
205 pgluglylylsAlaValTyrrleuLysleuAspleuLeuValAspGlyVal 222
66309 TGAGGGAAGGGCTGTCTACTGAAAGCTGAGCTGTGTGAGACGGTGTGC 66260
222 euAlaLeuArgrCysLeuGluGluPheSerAlaThrAlaAlaSerSerleu 238
66259 TGCCCTGGCGCTGCTGAGAAATCTCTCAACCACAGCAACACTCTCCT 66210
239 glyProGlnleuArgrleuCysGlnValSerGlyleuLeuAlaLeuArgr 255
66209 GGGCCCCAGACTCGTTGTGCGCAGGTGTCTGGGCTGTGGCGCTGGCGCC 66160
255 oclYserSerleuArgrIleArgrThrleuProTrrpAlaHisleuLysAla 272
66159 AGGGTCTTCCTTCGGATCGCACCTCCCTGGCTCATCTTAAGGCTG 66110
272 laProPheLeuThrTyrrPheGlyleuPheGlnValHis 284
66109 CCCCCTTCTTAACCTACTTTGGACTCTTTCAAGTTTCAAC 66072

seq_name: gb_hlcy:AC098923
seq_documentation_block:
LOCUS AC098923 179030 bp DNA linear HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-154B15, *** SEQUENCING IN PROGRESS
ACCESSION AC098923
VERSION AC098923.4 GI:17973852
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 179030)
Munry,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Day-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Deen,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flaeg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
COMMENT

```

```

Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
Hollins,B., Homel,F., Howard,S., Huber,J., Huiyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korah,J.,
Kovar,C., Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtenarge,O., Lieu,C., Liu,J., Liu,W.,
Loulisege,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabab,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogulu,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N.,
Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellrod,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wlaczek,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 179030)
Direct Submission
Worley,K.C.
Submitted (06-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17064677.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G10K
Center clone name: CH230-154B15
----- Summary Statistics
Assembly program: Phrap; version 0.990329p1r1st call to
findPhrapList
Consensus quality: 137096 bases at least Q40
Consensus quality: 145080 bases at least Q30
Consensus quality: 152950 bases at least Q20
Estimated insert size: 138980; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-1p estimation
Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a working draft sequence. It currently
* consists of 68 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 7467: contig of 7467 bp in length
* 7468 7567: gap of unknown length
* 7568 12115: contig of 4548 bp in length
* 12116 12215: gap of unknown length
* 12216 20313: contig of 8098 bp in length
* 20314 20413: gap of unknown length
* 20414 25589: contig of 5176 bp in length
* 25590 25689: gap of unknown length
* 25690 31434: contig of 5745 bp in length
* 31435 31535: gap of unknown length
* 31535 37595: contig of 6061 bp in length
* 37595 37596: gap of unknown length

```

```
* 37696 42172: contig of 4477 bp in length
* 42173 42272: gap of unknown length
* 42273 46234: contig of 3962 bp in length
* 46335 46334: gap of unknown length
* 50605 50705: contig of 4271 bp in length
* 50706 55202: gap of unknown length
* 55202 55302: contig of 4497 bp in length
* 55302 55303: gap of unknown length
* 55303 58407: contig of 3105 bp in length
* 58408 61286: gap of unknown length
* 61287 61386: gap of unknown length
* 61387 64567: contig of 3081 bp in length
* 64568 64567: gap of unknown length
* 64568 68498: contig of 3931 bp in length
* 68499 68598: gap of unknown length
* 68599 73449: gap of unknown length
* 73450 73549: gap of unknown length
* 73550 77402: contig of 3853 bp in length
* 77403 77502: gap of unknown length
* 77503 81027: contig of 3525 bp in length
* 81028 81127: gap of unknown length
* 81128 84144: contig of 3017 bp in length
* 84145 84244: gap of unknown length
* 84245 86680: contig of 2436 bp in length
* 86681 86780: gap of unknown length
* 86781 89210: contig of 2430 bp in length
* 89211 89310: gap of unknown length
* 89311 92339: contig of 3029 bp in length
* 92340 92439: gap of unknown length
* 92440 94951: contig of 2512 bp in length
* 94952 95051: gap of unknown length
* 95052 98586: contig of 3535 bp in length
* 98587 98686: gap of unknown length
* 98687 100597: contig of 1911 bp in length
* 100598 100697: gap of unknown length
* 100698 102766: contig of 2069 bp in length
* 102767 102866: gap of unknown length
* 102867 105237: contig of 2371 bp in length
* 105238 105337: gap of unknown length
* 105338 107305: contig of 1968 bp in length
* 107306 107405: gap of unknown length
* 107406 109157: contig of 1652 bp in length
* 109058 109157: gap of unknown length
* 109158 111554: contig of 2297 bp in length
* 111455 111554: gap of unknown length
* 111555 113704: contig of 2150 bp in length
* 113705 113804: gap of unknown length
* 113805 115933: contig of 2129 bp in length
* 115934 116033: gap of unknown length
* 116034 118193: contig of 2160 bp in length
* 118194 118293: gap of unknown length
* 118294 120414: contig of 2121 bp in length
* 120415 120514: gap of unknown length
* 120515 122311: contig of 1797 bp in length
* 122312 122411: gap of unknown length
* 122412 125628: contig of 3217 bp in length
* 125629 125728: gap of unknown length
* 125729 128545: contig of 2817 bp in length
* 128546 128645: gap of unknown length
* 128646 130563: contig of 1918 bp in length
* 130564 130663: gap of unknown length
* 130664 132945: contig of 2282 bp in length
* 132946 133045: gap of unknown length
* 133046 134556: contig of 1511 bp in length
* 134557 134656: gap of unknown length
* 134657 136478: contig of 1822 bp in length
* 136479 136578: gap of unknown length
* 136579 138159: contig of 1581 bp in length
* 138160 138259: gap of unknown length
* 138260 140456: contig of 2197 bp in length
* 140457 140556: gap of unknown length
* 140557 142544: contig of 1988 bp in length
```

```
* 142545 142644: gap of unknown length
* 142645 143784: contig of 1140 bp in length
* 143785 143785: gap of unknown length
* 143785 143884: gap of unknown length
* 143885 145304: contig of 1420 bp in length
* 145305 145305: gap of unknown length
* 145405 145404: gap of unknown length
* 147391 147391: contig of 1987 bp in length
* 147392 147392: gap of unknown length
* 147492 148908: contig of 1417 bp in length
* 148909 149008: gap of unknown length
* 149009 150147: contig of 1139 bp in length
* 150148 150247: gap of unknown length
* 150248 151752: contig of 1505 bp in length
* 151753 151753: gap of unknown length
* 151853 153441: contig of 1589 bp in length
* 153442 153442: gap of unknown length
* 153542 154953: contig of 1412 bp in length
* 154954 155053: gap of unknown length
* 155054 156244: contig of 1191 bp in length
* 156245 156344: gap of unknown length
* 156345 158095: contig of 1751 bp in length
* 158096 158195: gap of unknown length
* 158196 159309: contig of 1114 bp in length
* 159310 159409: gap of unknown length
* 159410 160682: contig of 1273 bp in length
* 160683 160782: gap of unknown length
* 160783 162602: contig of 1820 bp in length
* 162603 162702: gap of unknown length
```

```
alignment_scores:
  Quality: 385.50      Length: 176
  Ratio: 3.779        Gaps: 2
Percent Similarity: 57.955 Percent Identity: 51.136
```

alignment_block:

US-09-245-198a-4 x AC098923/rev ..

Align seg 1/1 to reverse of: AC098923 from: 1 to: 179030

```
160 GYVALASPGLYThValSerGIYTPGLUGLUALAaRgIIeAsnSense 176
|||||
26419 GGTGTGATGGACAGTGTGAGTGGAGAGACCAAAATCAACAGCTC 26370
176 rserProLeuArqTYrAsnArGInIIeGIYLPheIIeValTrpArA 193
|||||
26369 CAGCCCTCGCCGCTATGACCCGACATTTGGGAATTTACGGTCATCAGG 26320
193 IaGIYLeuTYrTYrLeuTYrCys..... 200
|||||
26319 CTGGGCTCTACACTGCTACTGTACGTAAAGCCTGGCCTCCATGGTAg 26270
200 ..... 200
26269 AGCGATGCTAAGGGAGAGAGGCTTGGCAAGAAATGGCTGGAGTGGGA 26220
200 ..... 200
26219 GAACCTGGGTTTCATGAGAGAAAGATCCCTGATTTTCATGAGCGACGCA 26170
201 .....GlnValHisPheAspGlu 206
|||||
26169 GAGCTGTGATTTGCTGTCTCTCTGTCACGCCAGTCACACTTATGATAG 26120
207 GYLYSAlaValTYrLeuLysLeuAspLeuValaSPGLYValLeuAl 223
|||||
26119 GGGAGGACACTTACTTAAGCTGAGCTTGCTGGTGAATGGTGTCTGGC 26070
223 aLeuArqCysLeuGIUGLUPheSerAlaTrhAlaIaSerSerLeuGlyP 240
|||||
26069 CCTGACACTGCTGGAATAATATATCAGCCACAGTACAGAGATCTCTGGGC 26020
240 rGInLeuArqLeuCysGlnVal..SerGIYLeuLeuAlaLeuArqProGI 256
|||||
26019 CCAGATCCGTTTGGCGCCAGGCCGCTGGCGCTGTTGATCCGCGGCCGG 25970
```

256 ySerSerLeuArgIleArgThIleuProTrrAlaNIleuLySAlaAlaIAP 273
|||||
25969 GCTTCACTTTCAGACCCGACACCCCTTAACCCCTTAACACCCCTCCG 25920
273 roPheLeuThyTrpPheGlyLeuPhe 281
|||||
25919 CCGTCTATCCCGCCCGGACGACTCTT 25894
seq_name: gb_pl:Ap002482
seq_documentation_block:
LOCUS Ap002482 187835 bp DNA linear PLN 10-JUN-2000
DEFINITION Oryza sativa genomic DNA, chromosome 1, clone:p0706B05.
ACCESSION Ap002482
VERSION Ap002482.1 GI:8468009
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (cultivar:Nipponbare) DNA, clone:p0706B05.
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 187835)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:p0706B05
Published only in Database (2000) In press
2 (bases 1 to 187835)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (07-JUN-2000) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abrr.affrc.go.jp,
URL:http://www.dna.affrc.go.jp:82/, Tel:81-298-38-7441,
Fax:81-298-38-7468)
The orientation of the sequence is from T7 to SP6 of the PAC clone.
Genes were predicted from the integrated results of the
following:GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as
SplicePredictor (October 1998 version). The genomic sequence was
searched against the non-redundant database NRP (PIR, SWISSPROT,
GENPEP, PDB) from MAF DNA bank and the cDNA sequence database at
RGP. Protein similarities of the coding regions were searched
against NRP with BLASTP2.0. ESTs represent the identified cDNA
sequences using BLASTN2.0 with the corresponding DBJ accession no.
and RGP clone ID.
Detailed information on overlap and assembly quality together with
annotation of this entry at
http://www.dna.affrc.go.jp:82/genomicdata/Genomefinished.html.
Location/Qualifiers
1. 187835
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="p0706B05"
1765..3195
/note="Similar to Streptomyces clavuligerus isopenicillin
epimerase (P18549)."
/codon_start=1
/protein_id="BAA96610.1"
/db_xref="GI:8468010"
/translation="MASLQSGDAANAGVDADVDGAASPPSAPKRRACAGAAITDAE
VRAPEAHDRRAVRLNNGTFCGCPASVYAAARAGRLFLSDPAFYPHILQPGARSR
AAVAAGACDASEVSLVDNTTAATIMOVANSFAGDFARGDVLMFLYTGCSIK
NSHAYAKAGATVEVPLRPVSPDVAIEFRALAVARDGGRRRLAVIDHTTA
MPVALIPRKELVAICREGVDFEVDAHAAGVVDVDRIGADFYASNLKWFCCPS
AVAFIPIRKDDPVSKLHPVSSSEYNGSLMESAMIGRDYSQLVVDPVDVFNRP
DGVGEGIRRRHDKVEMGTMLAAAGTFLCTPPEMCSMLVGLPGSLGVGSEDDAV
GLRTMLRKQFKEVPLVLYNSKAAADAPPEVKNQNDPRTGYRISHQYVNVREYE
ALRDVAKLVADGTCRKLREPKEVH"
join(6984..7174,7451..7576,7735..7846)
/note="hypothetical protein"
CDS
source
FEATURES
CDS
1..187835
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="p0706B05"
1765..3195
/note="Similar to Streptomyces clavuligerus isopenicillin
epimerase (P18549)."
/codon_start=1
/protein_id="BAA96610.1"
/db_xref="GI:8468010"
/translation="MASLQSGDAANAGVDADVDGAASPPSAPKRRACAGAAITDAE
VRAPEAHDRRAVRLNNGTFCGCPASVYAAARAGRLFLSDPAFYPHILQPGARSR
AAVAAGACDASEVSLVDNTTAATIMOVANSFAGDFARGDVLMFLYTGCSIK
NSHAYAKAGATVEVPLRPVSPDVAIEFRALAVARDGGRRRLAVIDHTTA
MPVALIPRKELVAICREGVDFEVDAHAAGVVDVDRIGADFYASNLKWFCCPS
AVAFIPIRKDDPVSKLHPVSSSEYNGSLMESAMIGRDYSQLVVDPVDVFNRP
DGVGEGIRRRHDKVEMGTMLAAAGTFLCTPPEMCSMLVGLPGSLGVGSEDDAV
GLRTMLRKQFKEVPLVLYNSKAAADAPPEVKNQNDPRTGYRISHQYVNVREYE
ALRDVAKLVADGTCRKLREPKEVH"
join(6984..7174,7451..7576,7735..7846)
/note="hypothetical protein"
CDS
CDS
/codon_start=1
/protein_id="BAA96611.1"
/db_xref="GI:8468011"
/translation="MVVFFPAHPIHGTHHNNSTPGVATEGGGSGGKSMADAE
IDRRRRARRRRRPREIADGTLDPSQASRLYSLGFGSNFNPQILEKTYTQLCAFSQNA
KVAREASRFSTIAVMATIDCAEDQSKCFVLYTKHLFFWT"
complement(join(8422..8775,8875..9168,9253..9516,
9642..9908,9986..10830,10934..11207,11317..11842,
12632..12870,12993..13214,13312..13487,13585..13981))
/note="Similar to Arabidopsis thaliana chromosome 2, BAC
F14M4; putative ABC transporter (AC004411)"
/codon_start=1
/protein_id="BAA96612.1"
/db_xref="GI:8468012"
/translation="MPESMRDAEANASSASVANAADSPGNGKGGGGAAGAAR
AASASARVDPFHLFAFDKTDALAEFTLGLVANGALPFEVTLFGULIAFGA
MGHIDVYNRSMVSEFTYLAASAVALPVOYTCWMTTGERQARINVLKTLTRE
IAFDKTYNTGEVGRMSGDTVLIDQAMGEVGFIDLVTFGLGFTVAHQGLLTL
VMAATIPPLVAVGAVMSNVVAKMASLQAAVAEESVVEQTLGIRVASFTGKQAV
EKYNSLSKAVKSGVREGLAAGLGMGVVLLFCGYSLGIWYGAKLILGLYGAQVM
NVIRAVLGSIALGQASPPMKAFAGCGAAAYKMEYINRRPEIDAYSTGCMKPDIDG
DIERFDVYFSPTRPDEQIPRGSLSTPGTYVALNGSGSKSTVSLIERFDPL
GDVLIDVNLKEPOLNRIRKICGVSEPVLPAASTENATAYGRDNATDDEIRAAAL
ANASKFDKMPQGLDTSVGEHTQLDSGQKORAIARAILDPRILLDLDTASLDE
SERIVQALDRVMNRTTYVAHRLSTVRNADTAVIHQGLVEKGFHBLDPEGA
YSOLIKLOEANROKSDRKDPSGARSKQSLINOSASRSRSHSDFSVFEGMP
LGIDIOGSDNLCDDGPDVPLSRSLASLNKPEIPLVILGSIASVIGTFEFAILL
SNVKAFFPEPHILRKDSQFMSMEIFGAVYFPLSPVSSYLSIAGCRILIKRILMT
FEKYVNMETEFHPENSSGATGRLSADAKVAGLYGDALQLYOVQNTLLGLVLA
FVSNWEJLILILPLILGLNGWIDMFIDGSADAKMIEASQVANDAVSIRTV
SFSAEKVMDLYKRCGPLETGIIRTIISIGIGVAFLLFVYVASFAGAVLVE
NKTTFPFRVFLALMAAIGVSSQSTLIDSSKAKSAVSIFAIVRKSIDPSEDA
GVYETLHNGIEFQHVSEFRYPTRPDVIFRDLCTIHSGKATLVAGSGSKSTAISL
LQRFYDDVGHILIDVGDIOKFLKMLROOMGLVNDPDALENDVRANIVGKGEAT
ESRTIEMAKLANAKFTISSHQGCTTYGEGNQLDSGQKORIAIARILKDPILL
DEATSAIDAESERVQDALDRVMNRTTYVAHRLSTIQNDLIAVYKNGVILTEKGN
DTLMNIDGAVASVALHSAAS"
join(16079..16164,17463..17646)
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA96613.1"
/db_xref="GI:8468013"
/translation="MAHTHRGACVRYSAAGRVSAQDSCAARCDVDAIGPCVT
SELFYSYRKSHKILKPCPNRGRDLPGNGLQGRREIYHEHNSN"
join(19104..19259,19684..19745,19879..20015,20087..20140,
21463..21598,24585..24645)
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA96614.1"
/db_xref="GI:8468014"
/translation="MOYCAVARYARVANVANSRROKPCGRSIDVTMLKTTNLSLP
LRPDETGVDPHGATASRGDRDRDRDYARAABHGCGRKREISRRRRRRVYRSIR
TQVYSHMIDLRTHWEGDYTYLHHEBEVYISGVGEVPTALDTFTTCPSFFLIYVR
ILSGDLIVERCHINSHSAPLVYIESIGRSEVNAHQFT"
complement(join(27285..27309,29006..29087,29375..29474,
29539..29601))
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA96615.1"
/db_xref="GI:8468015"
/translation="MTRAAVYHLNTSSSTAGCLAAMLGRGICQRMSRDQAASSIE
AMDLASGAGMALKKRGKREIKLKSSKCAPAEASRRMPTRM"
complement(30314..31296)
/note="3' LTR"
join(31975..32199,32416..32620,32656..33163,33648..33771)
/note="hypothetical protein"
/codon_start=1
/protein_id="BAA96616.1"
/db_xref="GI:8468016"
/translation="MPEDTVARAGESRPRPYPRSVHLLSFAMAGLIPPSREFHY
IDFETIHALHLAPNAVNTLAFVILCEMFIGRSTRAMLAVSDLGCEYDAVDRGLGN
QGLTGAVYFGDYFHCRIAPLOERSRGMEYTGPDPMKTHVDGGLAASGCHRRAND
SRMDPILCSDRDRESILAVMTVAGRGRRSGAGGCGDGTGSSGATAGCGRNGSS

CDS
GGDGSRAPGPNRPGGDSASDPKRRKMSFPPSPHGAERTADRPAGHKCPAA
SGVGRKKRLRKIGCTEPCRCGSFIEPPKWTFRPRPRDRPLSPSGTDESGSEGG
SDDMDVVGAGIGLGP"
complement(join(34469..3635),36628..37001))
/note="Similar to Oryza sativa GAG-FOL precursor
(AB030283)
internal stop codon:complement(36927..36929
probably inactive because stop codons and frameshift
positions are included in CDS"
/codon_start=1
/pseudo
34607-34613
/note="one bp frameshift insertion"
34922-34925
/note="one bp frameshift deletion"
35799-35794
/note="one bp frameshift deletion"
complement(40163..41147)
/note="5' LTR"
complement(join(41162..41301,41962..42181,44317..44345,
44535..44778,46639..46732,47583..47695))
/note="hypothetical protein"
/codon_start=1
/protein_id="BA096617.1"
/db_xref="GI:8468017"
/translation="MLTTTIDRVNRVHCRFIDLGEFPLATMHOQOVSIRKIRVDY
VARRPATICMDPFEFVGMILDMRRAAGSVISGTTATANERPAPCALPPM
TMRRRGVAPSPSTGCAATNDEEEEARASGATSDKNAASVAALIVTCRCILM
CSLMKASVITLGAAGSNIVFMVLSDSVLAHRLSDGTGASRESEFAASAIARL
HHHAAPCCLRPNQARETPADRRKSFLLSTGGCRPLGIDGKKMGHDLVCVLY
FTN"
join(48744..48975,49715..49749,49993..50176,51131..51258,
51368..51468,51518..51630,51771..51915,52454..52930,
53116..53324,53464..53712,53894..53976)
/note="hypothetical protein"
/codon_start=1
/protein_id="BA096618.1"
/db_xref="GI:8468018"
/translation="MLAGSDHHRAGPRGRGRCHRRVYGEKMRMPRDAREKROPHM
DEVVSLMPEMIVLEFVLAESSGRHARLRGRGLLEDNDCKSLVSGSOTMRKVN
GDITKNTSSTLPPPIRPDKIVFSSLMFGCPDEAMLEGBGLVSAMATNLALSTWA
RAKAVLLIGVSTIESIRSLSTLYSPFVPRSONVLAQRIDVANSRKNNKKTWIDH
SYTKNKKKMGAGGAGGVORPGDGLAAROGGMSGANCPCACANPLPPAA
HRRFSLAPVPADNALPLSLPVGGRALPLISRSVGRGVAAETRRRRRAGEG
RSGGGDRDRGRRLAVGRKKGHONGRLATGPATLSTVLAHLSAGHDAPASVL
PRPRCKGRRLATATRVGITSKMAASGGGFCGCGGAGGSRGAGCGGAADG
KRRRAPANDVYTRVIGICMHTAMSSSPYVSIVSYFVGNIRCKEEOXVPEAA
GADGVADRLLAAAPELSTFAKTKETTVAAAGAAARSLKQPRHPHRLSGFEES
AADDISMPAVGCGMMWCCRRNGERDMDSLPLPPLKPLAAIASALARSLPADG
ROYRHISATSAKPSKTAEGVKLHRFO"
complement(54260..54721)
/note="ESTs AU031727(R0930),D24036(R0930) correspond to a
region of the predicted gene.
hypothetical protein"
/codon_start=1
/protein_id="BA096619.1"
/db_xref="GI:8468019"
/translation="MAGAAMPAPAGAGNGGFCNCFPYSVYAESGAVBAADVAG
VYGMIREAKARLPADPARAVAVDFREDPELSTDSLSVSWTRIGLEADYGMKPP
SHVIPPAYTPFAVAALIGAPVPVKGTGARIMTCCVEDDHLPAKKEELKADK"
complement(57466..57918)
/note="EST AU032068(R3431) corresponds to a region of the
predicted gene.
Similar to Nicotiana tabacum mRNA for HSR201 protein.
hypersensitivity-related gene (X95343)"
/codon_start=1

alignment_scores:
Quality: 145.50 Length: 314
Ratio: 0.983 Gaps: 15
Percent Similarity: 47.134 Percent Identity: 26.752

alignment_block:

us-09-245-198a-4 x Ap002482 ..
Align seg 1/1 to: Ap002482 from: 1 to: 187835
13 LeuProLeuPro...ArgSer.LeuGlySerArgAspGlyAlaValA 28
|||||
32532 CTTCATTTGCCGATCGCACCCCTCCAGAGACGGCTCCGTCGCCATGCG 32581
28 rglInlaInProProlaProMetaIaAlaIaArgSerGlaArg 44
:::|
32582 AGTACACTGGGCCACAGATCCATG.....CGCACCATGTGGGTGAG 32625
45 ArgGlyArgArgGlyGluProGlyThrAlaLeuLeuValProLeuAlaLe 61
|||||
32626 CGCTGGAGCTAGGGCGAGAGAGGCGCGAAGATGATGATCTGCCAGTCT 32675
61 uGlyLeuGlyLeuAlaLeuAlaCysLeu..... 70
|||||
32676 GGGTCTGTGACATCGCCGAGCAACACATGATTCAGATGATCTCTCCCT 32725
70 70
32726 CTGACGCGACCGGACCGGAGAGATTCCTGCGCTGATGACGTCCTCG 32775
71GlyLeuLeuLeuAlaValSerle 79
32776 GCGCTGCGAGGGGTCGCTCCCGCGGAGCGCGCTGCGGTGTCGAT 32825
79 u.....GlySerArgAlas 84
32826 GGCACGGGACACAGTGTGCGACCCGACGTCGCGGATGCGGCTC 32875
84 erLeuSerAlaIngluProAlaIngluIngluLeuValAlaIngluInasp 100
32876 CAGCGGGGCGACGCTCCAGGCGCCGCTCTAATCTG..... 32916
101 GlnAspProSerGluLeuAsnProGlnThrGluInuSerGlnAspProAl 117
32917 ..GCCCGGGGAGATTCGCCGACGACCGGAAAGGAAACGAAATG 32963
117 aProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyA 134
32964 TCCAGATTTGGGCGCCCTTCACCCACCATGAGAGGGGCGCGAGCGAAC 33013
134 rGLVSThrArgAlaArgArgAlaIleAlaIaHisTyrGluValHisPro 150
33014 GCGGACCGACCCACCGGGGCGCCACAAATGCC.....CTGCGG 33051
151 ArgProGlyGlnAspGlyAlaGlnAlaGlyValAsp..... 162
33052 CGTCCGGGCTCGGACGGAAGAACGCGCTCCGGAAGATAGGCAACG 33101
163GlyThrValSerGlyTyrPgluGluAlaArgIleAsnSers 176
33102 GAGCGCTGCCGGGGAGTTTCATCGACCCCAAGTGA...CCTCAA 33148
176 erSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArg 192
33149 AGCGCCCTC.....CGAGTTGATTCGCAGATCCAA 33183
193 AlaGly...LeuTyrTyrLeuTyrCysGlnValHisPheAspGluGly 208
33184 CCGGCTGCTTCATCCACTCGGTTTCCTTTTTCCTTCCTTCCTTT 33233
208 sAlaValTyrLeuLysLeu.....AspLeuLeu.....ValA 219
33234 CTCTCTTTTAACTTGATCGGGCTCTCTAGTGAATTCATCCATCC 33283
219 spGlyValLeuAlaLeuArgCysLeuGluInuPheSerAlaThrAlaAla 235
33284 CAGGGGCTCTCGACGTCGATCCCTCACAGGGGTGAATTCGAGCATC 33333
236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAl 252


```

:::|||||:::|||||:::|||||
33334 GGATCGCGCTTGTCGCGCTGATCAGATCCAT.....GGAGTCTCTCCG 33377
252 aLeuArg.....ProGlySerSerLeuArgIleArg 262
|  |||:::|||||
33378 AGCATCTTGCGGGGCCACGCGCCGCGGCTCCGG 33417

```

seq_name: gb_ba:SC1C2

seq_documentation_block:

LOCUS SC1C2 42210 bp DNA linear BCT 15-JAN-1999
 DEFINITION Streptomyces coelicolor cosmid 1C2.
 ACCESSION AL031124
 VERSION AL031124.1 GI:3355667
 KEYWORDS 3-isopropylmalate dehydrogenase large subunit; 3-isopropylmalate dehydrogenase small subunit; 3-isopropylmalate dehydrogenase; branched-chain amino acid aminotransferase; carboxyl transferase; delta-1-pyrroline-5-carboxylate dehydrogenase; gltX; glutamyl-tRNA synthetase; histone-like DNA binding protein; hydrolase; lIve; leuB; leuC; leuD; lyase; secreted lyase; transfer-RNA-Gln; transfer-RNA-Glu; ureaB; urease alpha subunit; urease beta and gamma subunits; ureC.

SOURCE
 ORGANISM Streptomyces coelicolor A3(2).

REFERENCE
 AUTHORS Streptomyces coelicolor A3(2).
 JOURNAL Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 REFERENCE 1 (bases 1 to 42210) Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 AUTHORS Murphy, L. and Harris, D.
 JOURNAL Unpublished

REFERENCE
 AUTHORS 2 (bases 1 to 42210) Parkhill, J., Barrell, B. G. and Randalream, M. A.
 JOURNAL Direct Submission
 REFERENCE 3 (bases 1 to 42210) Submitted (22-JUL-1998) Streptomyces coelicolor sequencing project,
 Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
 REFERENCE 4 (bases 1 to 42210) CB10 ISA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
 David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
 Colney, Norwich, Norfolk NR4 7UH, UK

REFERENCE
 AUTHORS 3 (bases 1 to 42210) Redenbach, M., Kleiser, H. M., Denapilte, D., Eichner, A., Cullum, J.,
 Kinashi, H. and Hopwood, D. A.
 JOURNAL A set of ordered cosmids and a detailed genetic and physical map
 for the 8 Mb Streptomyces coelicolor A3(2) chromosome
 REFERENCE 4 (bases 1 to 42210) Mol. Microbiol. 21 (1), 77-96 (1996)

JOURNAL
 MEDLINE
 COMMENT

Notes:
 Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC.

Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.

(URL: http://www.sanger.ac.uk/Projects/S_coelicolor/) CDS are numbered using the following system eg SC787.01c. SC (S. coelicolor), 787 (cosmid name), .01 (first CDS), c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov/jp/jun/cgi-bin/frameplot.pl>. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an upstream codon (atg, gty, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence

overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 1C2 lies between 8D9 and 7A1 on the AseI-B genomic restriction fragment.
 Location/Qualifiers

```

source
1..42210
  /organism="Streptomyces coelicolor A3(2)"
  /strain="A3(2)"
  /db_xref="taxon:100226"
  /clone="cosmid 1C2"
  1..1228
    /gene="SC1C2.01"
    /gene="SC1C2.01"
    <1..1228
      /gene="SC1C2.01"
      /note="SC1C2.01, probable delta-1-pyrroline-5-carboxylate dehydrogenase, partial CDS. len >408 aa; similar to many e.g. TR:050443 (EMBL:AL010186)
      delta-1-pyrroline-5-carboxylate dehydrogenase (M. tuberculosis) (543 aa), fasta scores: opt: 1730 z-score: 2214.2 E(): 0, 65.1% identity in 407 aa overlap, and
      PUN2_HUMAN delta-1-pyrroline-5-carboxylate dehydrogenase (563 aa), fasta scores: opt: 1273 z-score: 1730.4 E(): 0, 48.5% identity in 408 aa overlap. Contains P500687 and P500070 Aldehyde dehydrogenases glutamic acid active site and cysteine active site and Pfam match to entry PF00171 aldedh. Aldehyde dehydrogenases, score 114.70, E-value 1.8e-30"
      1..8e-30"
        /codon_start=2
        /transl_table=11
        /product="delta-1-pyrroline-5-carboxylate dehydrogenase"
        /protein_id="CAI19968.1"
        /db_xref="GI:3355668"
        /db_xref="SP|P086502"
        /translation="IDSPCELDIFMFRNVHVARNIILAEOPANSQGVNRRHDPLEGG
        FYVAIRPFNSALIAANLPTAPALMGVNVKPSPTOTHAVALMLQLEEGLPKGVN
        LMGDIAVSEVLAHEHDLAGIFHTGTFKHLMKGVNNIKRYRTPVLTGETGK
        DLEVHNSARAVYKTRLTGATGATGYSQKSSATSRATTPASINDDGKEFAAEVTL
        TMGDVTLDSNFGVAVIDERSFARKKAIDRAKDECTIVAGSDYDSVYFVRPTV
        ECDPENREYRTEYFEPFLAVHYVDSADADYAMLIQMSVSDYALATGVSINDRA
        AATYMEKLRVYAGNFYINDKSTGAVVGQPFGGGRASGTNDKAGAPQNLKRWTLTRAI
        KETLVAPDTYTYPMG"
        1..3807
          /note="true overlap with cosmid 8D9"
          2..1189
            /gene="SC1C2.01"
            /note="Pfam match to entry PF00171 aldedh, Aldehyde dehydrogenases, score 114.70, E-value 1.8e-30"
            464..487
              /gene="SC1C2.01"
              /note="P500687 Aldehyde dehydrogenases glutamic acid active site"
              548..583
                /gene="SC1C2.01"
                /note="P500070 Aldehyde dehydrogenases cysteine active site"
                1246..1283
                  /note="hairpin_loop with 18bp stem"
                  1522..1527
                    /note="possible RBS upstream of SC1C2.02"
                    1537..1668
                      /gene="SC1C2.02"
                      1537..1668
                        /gene="SC1C2.02"
                        /note="SC1C2.02, questionable ORF, len: 43 aa"
                        /codon_start=1
                        /transl_table=11
                        /product="hypothetical protein SC1C2.02"
                        /protein_id="CAI19969.1"
                        /db_xref="GI:3355669"
                        /db_xref="SP|P086503"
                        /translation="MAETTVRRVRRLHRTSDSDRKNAALQSLDRDNGATGCH"
                        1744..2787
                          /gene="leuB"
                          1744..2787

```



```

/gene="leub"
/note="SCIC2.03, leub, probable 3-isopropylmalate
dehydrogenase, len: 347 aa: similar to e.g. LBD3 CORCL
3-isopropylmalate dehydrogenase (EC 1.1.1.85) (340 aa),
fasta scores; opt: 1450 z-score: 1318.1 E(): 0, 68.08
identity in 338 aa overlap. Contains PS00470 Isocitrate
and isopropylmalate dehydrogenases signature and Pfam
match to entry PF00180 isodh, isocitrate and
isopropylmalate dehydrogenases, score 356.60, E-value
1.3e-105"
/codon_start=1
/transl_table=11
/product="3-isopropylmalate dehydrogenase"
/protein_id="CA119970.1"
/db_xref="GI:3355670"
/db_xref="SPTREMBL:O86504"
/translation="MSRLAVIPEDGIGQEVYAGKLVAVLPDYKLETKERDF
GARVHATGETLTDADLDALKAHDAIILGADPSVSGLEKGLIKLRFADHVN
LRPSKILPGVATPLAQPEIDVYVREGTEGPTGNGGTIRKGTREHEVATVNTAY
GVERVYVDAFARAQARPRKRLTLHKNNVLTFGHMTNFNKAENPEVTDYLAH
DAATIFLVDPARFVIVTDNLEGGDIITDLAAVSGGIVAAAGNINPSGDFMSEP
VHGSAPIAGCGKADPTATVLSVALLRLRHGTDEARIEDAVSADLGERGDLPARST
SEIGDTLAARVAG"
/length=159
/gene="leub"
/note="Pfam match to entry PF00180 isodh, Isocitrate and
isopropylmalate dehydrogenases, score 356.60, E-value
1.3e-105"
/length=2458
/gene="leub"
/note="PS00470 Isocitrate and isopropylmalate
dehydrogenases signature"
/length=3038
/note="possible RBS upstream of ilvE"
/length=3052
/gene="ilvE"
/length=3052
/note="SCIC2.04, ilvE, probable branched-chain amino acid
aminotransferase, len: 362 aa: similar to many, e.g.
ilvE_BACSU putative branched-chain amino acid
aminotransferase (362 aa), fasta scores; opt: 1047
z-score: 1401.5 E(): 0, 44.7% identity in 360 aa overlap.
Contains PS00770 Aminotransferases class-IV signature"
/codon_start=1
/transl_table=11
/product="branched-chain amino acid aminotransferase"
/protein_id="CA119971.1"
/db_xref="GI:3355671"
/db_xref="SPTREMBL:O86505"
/translation="MTPTITELKPSAHLSDSPRAAILANPGGRHETDMMVITKTE
GRGMHGOQLVYPAPLSLDPRATWLIHTAQELFEBELAKYRRPDSVAIFRRKNGARPOA
SSRRIGMPELPVDTLEACDALVAODEKVPAPHGGESLYLRFMATATEGLCLAVRQA
EXLFIIVASPAQVYFPGVAPSVINWSEDRVAPGMDAKGVAASLLAQAAEA
AAGCDQVCLDAIERKVELGEMNLVYFGNKTIVPSLITGLSEVTRSLSTTVARD
LAYEAEGRVSVDOMORSENGTLTEVPACGTAAVITPGTVARAGAOOOSGETGE
VTORLDAILLDIORGTVADPHGMHTLA"
/length=3763
/gene="ilvE"
/note="PS00770 Aminotransferases class-IV signature"
/complement(4214, .5644)
/gene="SCIC2.05c"
/complement(4214, .5644)
/gene="SCIC2.05c"

alignment_scores:
Quality: 140.00 Length: 335
Ratio: 0.828 Gaps: 18
Percent Similarity: 50.448 Percent Identity: 28.358

alignment_block:
US-09-245-198a-4 x SCIC2/rev ..

```

```

Align seg 1/1 to reverse of: SCIC2 from: 1 to: 42210

9 SerAlaArgLeuProLeu.ProArgSerLeuLysSerArgAspGlyG 25
||||| ||| |||: |||||: ||||| |||
9771 TCCGTCGCCCGAGGCGGATGCCGTGCGAGACGAGGCGGCGTCAAGAG 9722
25 Ly...AlaValArgGlnAlaGlnProProAlaProMetaLAlaArgArg 40
|||: |||: |||: |||: |||: |||: |||: |||
9721 GCGCCGACGCCGAGCGGAGCGGAGCGGTCGGGTCGA.....CGCGC 9681
41 SerGlnArgArg.ArgGlyArgArgGlyGluProGlyThrAlaLeuLeu 57
|||: ||||| ||| ||||| ||| |||: |||
9680 GGAACTCGCCCGCGAGATGCCCTCGCGAGACGAGGCGGTGAGTCGGCG 9631
57 aLProLeuAlaLeuLysLeuGlyLeuAlaLeuAlaCysLeu....Gly 71
||||| ||||| ||||| ||||| ||||| |||
9630 TCCGAGGCGCCCTCATGGCGGCTGCGCGGCGGCTGTCGGGCGG 9581
72 LeuLeuLeuAlaVal..... 76
|||: |||: |||: |||: |||: |||: |||: |||
9580 GTTCTGTAGTGTCTCCACACCTCCAGCCACAGCGTCCAGTCGGGTCGC 9531
76 ..... 76
9530 GGTGCGCGTCCGGGACGTATACAGTCGACGTAGCGCGGCGGCTTCGGCG 9481
77 .....ValSerLeuGlySerArgAlaSer 84
CGCGTCCGCGCGCGGTCGAGCAGGCGGCGCGGCGGCGGCGGCGGCGGCGC 9431
85 LeuSerAlaGlnGluProAlaGlnGluLeuVal.....AlaGlnG 99
||||| ||| |||||: |||||: ||||| |||
9430 CTCGCTCCACATCCAGGTCGCGCAGCAGCATTCGTCCTTGAGCGCGAAGT 9381
99 uAspGlnAspProSerGlu.LeuAsnProGlnThrGluGlnSerGlnAsp 115
|||: |||: |||: |||: |||: |||: |||: |||
9380 AATACAGAGGTGCGCGCTGCTCATGCGACCTCCGCGCGCGAGC..... 9337
116 ProAla.ProPheLeuAsnArgLeuValArgProArgArgSerAlaProL 132
||| |||||: |||: ||||| |||||: ||||| |||
9336 ..GCCGATGATGAGCTTCTCCAGCCCGCGCTGGGATCATATCTCAT 9290
132 ySgLYLArgLysThrArgAlaArgArgAla.....ILAlaLAlaHisTry 146
|||: |||: |||: |||: |||: |||: |||: |||
9289 GCGCGCGCGAGACAGCTCTCGCGCGCGCGCGCGCTGCGCCACACAC 9240
147 GluValHisProArgProGlyGlnAspGlyAlaGlnAla..... 159
||| ||| ||||| ||| |||: |||
9239 CGGCATGTCACATCCCGGCTCTTCTAGTCTGCTGCTGCTGCTGCA 9190
160 GlyAlaAspGlyThrValSerGlyTTPGluGlnAlaArgIleAsnSer 176
|||: |||||: |||: |||||: |||: |||||: |||: |||
9189 TGCATGTGATGCCCCCGCACCCCGGAAGATCGAGCGCGCCACAGG 9140
176 er...SerProLeuArgTyrAsnArg.....GlnIleGlyIlePhe 188
||| ||||| ||| |||: |||: |||: |||: |||
9139 GTCACCGTCCTCGCGGACACAGCGCGGCGGAAGATCCCGCGCGAGTTC 9090
189 IleValThrArgAlaGlyLeuTyrTyrTyrLeuTyrCysGlnValHisPheAs 205
||| ||| ||| ||| ||| ||| ||| ||| |||
9089 GTCGTCGCGGTGCGC.....GAAGCGGCACAGGA 9061
205 pGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyVal 222
|||: |||: |||: |||: |||: |||: |||: |||
9060 CGACGCGCGCGTGCACACGATGTGTGATGTAGATGTAGTGGCCAC 9011
222 euAlaLeuArgCysLeuGluGluPheSerAla..... 232
||||| |||: |||: |||: |||: |||: |||: |||
9010 TCGCCTCTCTGCTCTTACAGACGTCGGGCGGCGGCGACCTCGACACCTC 8961
233 ..ThrAlaAlaSerLeuGlyProGlnLeuArgLeuCysGlnValSer 248
|||: ||||| |||: ||||| |||: ||||| |||: |||||
8960 CAGCGCGCGCGCGCTGCGTGGTCCGCGCGCGCGAGATCTCCAGGTACA 8911

```

248 r.....glyleuleuAlaLeuArgProGlySerSerLeuArgIleArg. 262
: ::::::::::: ||||| ::||| :::::
8910 GCGTGTGAGCGCTGTGTCGGGCGGCGGTCGCGTGTGTCACG 8861
263 ThrLeuProTPrAlaHisLeuysAlaAlaProPheLeuThrTyr 277
||| ::::: ||||| ||| :::
8860 ACGACGTCGCCGCCGCGAAGCCGCGACGATGTGACGTCG 8816
seq_name: gb_ba:AL646059
seq_documentation_block:
LOCUS AL646059 190050 bp DNA linear BCT 07-DEC-2001
DEFINITION Ralstonia solanacearum GM11000 chromosome, complete sequence;
segment 3/19.
ACCESSION AL646059 AL646052
VERSION AL646059.1 GI:17427391
KEYWORDS
SOURCE Ralstonia solanacearum.
ORGANISM Ralstonia solanacearum
Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
Ralstonia.
REFERENCE 1 (bases 1 to 190050)
AUTHORS Salanoubat,M., Genin,S., Artiguenave,F., Couzy,J., Mangenot,S.,
Ariat,M., Billault,A., Broctier,P., Camus,J.C., Cattoiico,L.,
Chandler,M., Choisme,N., Claudel-Renard,C., Cunnac,S., Demange,N.,
Gaspin,C., Lavie,M., Moisan,A., Robert,C., Saurin,M., Schlex,T.,
Siquier,P., Thebaudt,P., Whalen,M., Winkler,P., Levy,M.,
Weissenbach,J. and Boucher,C.A.
TITLE Genome sequence of the plant pathogen Ralstonia solanacearum
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 190050)
AUTHORS Boucher,C.A.
JOURNAL Direct Submission
COMMENT Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
Cremieux, CP5706, 91057 Evry Cedex, France. Laboratoire de Biologie
Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS,
BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean
Dausset-CEPH, 27 rue Juliette Dodu, 75010 Paris, France, IMCM CNRS
118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA
URV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,
Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27,
F31326 Castanet-Tolosan Cedex. Laboratoire de Genetique Cellulaire
INRA, BP27, F31326 Castanet-Tolosan Cedex
Christian.Boucher@toulouse.inra.fr
http://sequence.toulouse.inra.fr/R.solanacearum.html.
FEATURES
source location/Qualifiers
1. 190050
/organism="Ralstonia solanacearum"
/strain="GM11000"
/db_xref="taxon:305"
complement(163..750)
/gene="RSC0382"
/note="RSC0350"
complement(163..750)
/gene="RSC0382"
/function="miscellaneous; not classified regulator"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon-usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not experimental
/product="PROBABLE RESPONSE REGULATOR TRANSCRIPTION
REGULATOR PROTEIN"
protein_id="CAD13910.1"
/db_xref="GI:17427392"
/translation="MAPTRVLVDPDPAEQVNLIRAGLAAGFVGVVDADIML
PERTAQOPKIVIVASASARDTEHCVCVATQHPAPRVILFTDNDASRIKAAAGI
TATVVDGKPTPRVAVDVAIVAREHERALAEIDNKTQTLAEKRVLEBRAGLLMQM
QLSDEAEKRLRKMAADRNIKIVAEADRVIVDMAG"

gene 937..1638
/gene="RSC0383"
/note="RSC0351"
937..1638
/gene="RSC0383"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon-usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not experimental
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD13911.1"
/db_xref="GI:17427393"
/translation="MTDFVYTPPAVYTGIPYRGAAGEPFRVRYCVGNVAARREMS
DPDRPPEFFCKPADAVRYADGTGQFVPTERDCHYEIELVATIGTRIVAS
AAQHIYGAIGLMDTRDLQNAKKGPRWETSKAEDGAPIGIYPAQTVASPDKGA
ITLSVNDKAHQHGDLSDIWSVETIAYLSRLFELRPGDLIFGTGEGVPSIGDLM
VGRIDGVGELHVXV"
1642..2292
/gene="RSC0384"
/note="RSC0352"
1642..2292
/gene="RSC0384"
/EC.number="5.2.1.2"
/function="small molecule metabolism; degradation; amino
acids"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon-usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not experimental
/product="PROBABLE MALEYLACETOACETATE ISOMERASE
(GLUTATHIONE TRANSFERASE ZETA 1) PROTEIN"
/protein_id="CAD13912.1"
/db_xref="GI:17427394"
/translation="WSIKLTYNFRSSASFVRALAEIKLPYDAPVPHLLKGGQAPD
FVKLNPAALPVLCDDGTEVNLQSLAIVYEETPEPTLLGSAADRAHIRALALA
CEHPLDNLNPRYLKTKHTFNVGDDARDWVRHWRRLGFALLETRLQSPPTGCCVD
TPTLADCLNPQVFNKGRFDVAVEDYPTLARVFEBHCAQAPAFQAAPADPAASA"
2289..2927
/gene="RSC0385"
/note="RSC0353"
2289..2927
/gene="RSC0385"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : probable
Gene name confidence : hypothetical
predicted by Codon-usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not experimental
/product="PROBABLE TRANSMEMBRANE PROTEIN"
/protein_id="CAD13913.1"
/db_xref="GI:17427395"
/translation="MTGPGPESAPLIGALSFGSLIVATIGAQNAAYVLRQGLRRE
YVIGVLTICALCDMALTAIVAGMGCTLISAPMLTAVRNAGAARFLIAGGRVAMPANVAFAG
RGARFLQARNDRKASHQVLAALALSLNPHRYLIDPVVILGAIAGGRVAMPANVAFAG
GAMCASTLWFSLSLGEARLDEPVFARVAMRVLDALIGAVMAIALTLMLMG"
complement(2924..3910)
/gene="ftsY"
/note="RSC0386: RSC0354"
complement(2924..3910)
/gene="ftsY"
/function="cell processes; cell division"

[illegible]

| | |
|-----------|---|
| TITLE | Direct Submission |
| JOURNAL | Unpublished |
| REFERENCE | 2 (bases 1 to 233912) |
| AUTHORS | Worley,K.C. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (09-Jan-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA |
| COMMENT | On Jan 16, 2002 this sequence version replaced gi:18092692.

Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GNASH
Center clone name: CH230-192N7

Summary Statistics
Sequencing vector: Plasmid; M7789
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329First call to
findphraplist

Consensus quality: 22407 bases at least Q40
Consensus quality: 22763 bases at least Q30
Consensus quality: 229761 bases at least Q20
Estimated insert size: 217555; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 6.6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a "working draft" sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 63113: contig of 63113 bp in length
* 63114 63213: gap of unknown length
* 63214 114898: contig of 51685 bp in length
* 114899 114998: gap of unknown length
* 114999 141849: contig of 26651 bp in length
* 141850 141849: gap of unknown length
* 141950 160892: contig of 18943 bp in length
* 160893 160992: gap of unknown length
* 160993 176073: contig of 15081 bp in length
* 176074 176173: gap of unknown length
* 176174 192709: contig of 16536 bp in length
* 192710 192809: gap of unknown length
* 192810 201398: contig of 8559 bp in length
* 201399 201498: gap of unknown length
* 201499 205776: contig of 4278 bp in length
* 205777 205876: gap of unknown length
* 205877 210968: contig of 5092 bp in length
* 210969 211068: gap of unknown length
* 211069 214358: contig of 3280 bp in length
* 214359 214458: gap of unknown length
* 214459 216239: contig of 1781 bp in length
* 216240 216339: gap of unknown length
* 216340 218395: contig of 2056 bp in length
* 218396 218495: gap of unknown length
* 218496 220455: contig of 1960 bp in length
* 220456 220555: gap of unknown length
* 220556 221854: contig of 1239 bp in length
* 221855 221954: gap of unknown length
* 221955 223563: contig of 1609 bp in length
* 223564 223663: gap of unknown length
* 223664 225111: contig of 1448 bp in length
* 225112 225211: gap of unknown length
* 225212 226214: contig of 1003 bp in length
* 226215 226314: gap of unknown length |

| FEATURES | | source | |
|---|---|-------------------|------------------------------|
| * | 226315 | 227488: | contig of 1174 bp in length |
| * | 227489 | 227588: | gap of unknown length |
| * | 227859 | 228852: | contig of 1264 bp in length |
| * | 228853 | 229652: | gap of unknown length |
| * | 228933 | 229663: | contig of 1011 bp in length |
| * | 229644 | 230063: | gap of unknown length |
| * | 230064 | 231343: | contig of 1280 bp in length |
| * | 231344 | 231443: | gap of unknown length |
| * | 231444 | 232645: | contig of 1202 bp in length |
| * | 232646 | 232745: | gap of unknown length |
| * | 232746 | 233912: | contig of 1167 bp in length. |
| location/Qualifiers | | | |
| 1. 223912 | | | |
| /organism="Rattus norvegicus" | | | |
| /db_xref="taxon:10116" | | | |
| /chromosome="Rf1" | | | |
| /clone="CH230-192N7" | | | |
| BASE COUNT 58900 a 5745 c 56429 g 58611 t 2227 others | | | |
| ORIGIN | | | |
| Alignment_scores: | | | |
| Quality: | 136.50 | Length: | 314 |
| Ratio: | 0.916 | Gaps: | 24 |
| Percent Similarity: | 47.452 | Percent Identity: | 31.210 |
| alignment_block: | | | |
| US-09-245-198A-4 x AC105470/rev .. | | | |
| Align seg 1/1, to reverse of: AC105470 from: 1 to: 233912 | | | |
| 3 | LeuLeuAspPhgGluLeuSerAlaArgArgLeuProLeuProArgSerle 19 | | |
| 32662 | CTTCTCGAT.....AGCCTAGCGGCGCTGCCG...AGTAGAGCAT 326255 | | |
| 19 | uGlySerArgAsp.....GlyGlyAlaValArgGlnA 30 | | |
| 32624 | AGGGTCTGTGGACCCGGACAGCCCTAGGGGGTCCGAGTGGACCGCG 32575 | | |
| 30 | lAglnProProAlaPrometAlaAlaArgSerGlnArgArgArgly 46 | | |
| 32574 | CCAAGCCCGCGTCCGCTTCTGGCGGAGTAGAGGCTCGCGGCTCATCG 32525 | | |
| 47 | ArgArgGlyGln...ProGlyThrAlaLeuValProLeuAlaLeuG 62 | | |
| 32524 | CGGCGAGTGGACCGCGCGCG.....CTCGG 32499 | | |
| 62 | ylEu.GlyLeuAlaLeuAla...CysLeuGlyLeuLeuAlaValVal 77 | | |
| 32498 | GGCGTGGGGGAGCGTCCGGGAGTCCCGCGGCGCTGACGACCTCTGTC 32449 | | |
| 78 | SerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGlu 94 | | |
| 32448 | GCC.....GCTCGTGGCAGGAGCCGCTCGGCTGGCCGACCATCTCG 32405 | | |
| 94 | luEuValAlaGluGlnAspGlnAspProSerGluLeuAsnProGlnThr 110 | | |
| 32404 | AC.....GCGGTGAGAGCAG.....CCCATGAGAGCAGCGGCGGCAC 32364 | | |
| 111 | GluGluSerGlnAspProAlaProPheLeuAsnArgLeuValArgProA 127 | | |
| 32363 | GAAACGGACGACGAGCCGCCCGCC.....AAGCGAGAGCGCC 32326 | | |
| 127 | rgArgSerAlaProGlyGly...ArgLyThrArgAlaArgAlaAla 142 | | |
| 32325 | GGTCGAGCTAGCCCTGGGGGCAAGCGCCCGGACCTCGCGGGGGTCTG 32276 | | |
| 143 | AlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAl 159 | | |
| 32275 | GGGGG.....GACACAGCGCGCGCCGAGCGGAGAACGAGAC 32238 | | |
| 159 | agLyValAspGlyThrValSerGlyTyrProLnuAlaAlaGlyLeuAsnSers 176 | | |

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2002, 10:48:59 ; Search time 2447.42 Seconds
(without alignments)
7571.777 Million cell updates/sec

Title: US-09-245-198a-3
Perfect score: 1373
Sequence: 1 atcgtatgttgaacttga.....gacaaatgtataatg 1373

Scoring table: IDENTIFY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 674847542 residues
otal number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 765.8 | 55.8 | 834 | 10 | BI766766 603056866 |
| 2 | 743.4 | 54.1 | 777 | 10 | BI819200 603038614 |
| 3 | 701.4 | 51.1 | 731 | 10 | BI871711 603395825 |
| 4 | 688 | 50.1 | 828 | 10 | BI596681 603243254 |
| 5 | 677.4 | 49.3 | 728 | 10 | BI870393 603395641 |
| 6 | 584.4 | 42.6 | 609 | 10 | BI966060 1672904.x |
| 7 | 581.4 | 42.3 | 1027 | 12 | AF163779 AF163779 |
| 8 | 531.6 | 38.7 | 910 | 10 | BG110063 602279667 |
| 9 | 516.8 | 37.6 | 531 | 10 | BI824443 603038693 |
| 10 | 493.6 | 36.0 | 918 | 10 | BE577781 602092080 |
| 11 | 458.2 | 33.4 | 1033 | 11 | AK020909 Mus muscu |
| 12 | 454.4 | 33.1 | 561 | 9 | AW763237 ur70d09.y |
| 13 | 442.8 | 32.3 | 587 | 10 | BG686319 602638232 |
| 14 | 440.2 | 32.1 | 456 | 10 | BI966255 1672904.y |
| 15 | 436.8 | 31.8 | 440 | 10 | BM128059 1109b06.y |
| 16 | 431.6 | 31.4 | 538 | 10 | BF821434 MRI-R7003 |
| 17 | 404 | 29.4 | 436 | 12 | AO890280 HS_3188_B |

| | | | | | |
|----|-------|------|-----|----|----------|
| 18 | 399.4 | 29.1 | 785 | 10 | BI762908 |
| 19 | 394 | 28.7 | 413 | 9 | AI422296 |
| 20 | 389.2 | 28.3 | 698 | 10 | BI906850 |
| 21 | 374.2 | 27.3 | 416 | 9 | AI291866 |
| 22 | 372.6 | 27.1 | 584 | 9 | AW917574 |
| 23 | 365.4 | 26.6 | 422 | 10 | BI677255 |
| 24 | 365 | 26.6 | 399 | 9 | AI913541 |
| 25 | 363.4 | 26.5 | 367 | 12 | AO100365 |
| 26 | 359 | 26.1 | 474 | 10 | BI965174 |
| 27 | 358.4 | 26.1 | 407 | 9 | AI221985 |
| 28 | 357.6 | 26.0 | 894 | 10 | BI908274 |
| 29 | 352.8 | 25.7 | 409 | 9 | AM131279 |
| 30 | 347.2 | 25.3 | 397 | 9 | AI669243 |
| 31 | 342.6 | 25.0 | 710 | 10 | BE858726 |
| 32 | 339.8 | 24.7 | 372 | 10 | BI677256 |
| 33 | 327.8 | 23.9 | 650 | 10 | BG404836 |
| 34 | 327 | 23.8 | 346 | 10 | BE858822 |
| 35 | 327 | 23.8 | 471 | 9 | AA221610 |
| 36 | 326 | 23.7 | 337 | 9 | AM195034 |
| 37 | 326 | 23.7 | 345 | 10 | BF439993 |
| 38 | 326 | 23.7 | 352 | 9 | AM204512 |
| 39 | 326 | 23.7 | 364 | 9 | AM291620 |
| 40 | 318 | 23.2 | 329 | 9 | AI695776 |
| 41 | 318 | 23.2 | 340 | 10 | BF195436 |
| 42 | 318 | 23.2 | 542 | 10 | BF041509 |
| 43 | 317 | 23.1 | 346 | 9 | AI760777 |
| 44 | 315 | 22.9 | 346 | 10 | BF940141 |
| 45 | 314 | 22.9 | 351 | 10 | BG054914 |

ALIGNMENTS

RESULT 1
LOCUS BI766766 834 bp mRNA linear EST 25-SEP-2001
DEFINITION 603056866F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5',
mRNA sequence.
ACCESSION BI766766
VERSION BI766766.1 GI:15758344
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 834)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1517 row: C column: 18
High quality sequence stop: 772.
Location/Qualifiers
1. 834
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5206217"
/clone_id="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size

[illegible]

Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shuraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1722 row: k column: 13
 High quality sequence stop: 776.

FEATURES
 source Location/Qualifiers

1..828
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5285892"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptPR (modified
 pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized to 10⁷ 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 155 a 278 c 223 g 172 t
 ORIGIN

Query Match 50.1%; Score 688; DB 10; Length 828;
 Best Local Similarity 95.3%; Pred. No. 4.6e-120;
 Matches 732; Conservative 0; Mismatches 30; Indels 6; Gaps 2;

QY 312 ggaactgaatcccccagacagaagaagccagatcctcgccttctcctgaacccagactagt 371
 DB 43 GGAATGGAATCCCGACAGACAGAAAGCCAGATCTCTGCGCTTCTGCAACGACTAGT 102
 QY 372 tcggcctcgcgaagatgcacttaaggccgaaacacgggctcgaagagcgatcgaagc 431
 DB 103 TCGGCTCGGAGAGTGCACCTAAAGCCCGAAACACGGGCTCGAAGACGATCGAGC 162
 QY 432 ccaatgaagatccatccagacacgagcagcagcagcagcagcagcagcagcagcagc 491
 DB 163 CCAATTATGAAGTTCATCCAGACACTGACAGACGACGACGACGACGACGACGAC 222
 QY 492 agtgaatgcctggaggaagccagaatcaacagcctccagcctctgcgtacacaacgcga 551
 DB 223 ATTGAGTGGCTGGAGAGAACCCAGAAATCAACAGCTCCAGCCCTCTGCGCTAACAAACGCCA 282
 QY 552 gatcgggagatctatagtcacccgggctggcctactactactactactactactactact 611
 DB 283 GATGCGGAGATTTATAGTCACCGGGGCTGGCTCTACTACTCTACTACTACTACTACTACT 342
 QY 612 tgaagaagggaagcctgctacactgaagcctgaagcctgctgctgaatgctgtctgagccct 671
 DB 343 TGATGAGGGGGAAGGCTCTACTGAAAGCTGAGACTTCTGCTGATGATGCTGCTGCGCCT 402
 QY 672 ggcctgcctggaggaatctcagccactgcgcgcagcttcctcctgcggccacagctcgccct 731
 DB 403 GCGCTGCTGGAGGAATTTCTACAGCCACTGCGGCGCAGTTCCCTCGGGGCCACAGCTCGCCCT 462
 QY 732 ctgcagaggtctgctggctgtgtgctcctgcggcagaggtctcctcctgcggatccgaacct 791
 DB 463 CTGCGCAGGCTGTGGCTGTGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 522
 QY 792 cccctggcccatcgaagcctgcctcctcctcctcctcctcctcctcctcctcctcctcctcct 851
 DB 523 CCGCTGGGCGCATCTCAAGGCTGCGCCCTTCTCTACTACTTGTGAGACTCTTCCAGAGTTCA 562
 QY 852 ctgaggggcccctgctcctcccaagctgctccaggtctgcgcgtccctcctcctcctcctcctc 911

|||||
 DB 583 CTGAGGAGGCCCTGCTGCTCCCGCAGTCGCCAGGCTGCGGCTCCCTCGACAGCTTC 642
 QY 912 tggagcccccgtccctctgcgcacccctcagcagcgtcttctgctcagagccctgcctcc 971
 DB 643 TGGGCAACCGGCTCCCTCTGGCCACCCCTCAGCAGCTTTGCTCCAGACCTGCCCCCTCC 702
 QY 972 ctctagaagctgcctgctgcctgtctcagcgtgtctcctcctcctcctcctcctcctcctc 1025
 DB 703 CTCTAAGAGGCTGCTGCGGCGCTGTTCACAGTTGATTCATCCGCCACATTAATACCGTTA 762
 QY 1026 attccacactctatcttaaacctcccccacgcgcacctcctcctcctcctcctcctcctc 1073
 DB 763 TTTCCCACTCTTATCTTACAACTTCCCGCCACAGACCAATCATTCACA 810

RESULT 5
 BI870393
 LOCUS BI870393 728 bp mRNA linear EST 11-OCT-2001
 DEFINITION 60395641F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405459 5',
 mRNA sequence.
 ACCESSION BI870393
 VERSION BI870393.1 GI:16044066
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 728)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12034 row: a column: 12
 High quality sequence stop: 728.

FEATURES
 source Location/Qualifiers

1..728
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5405459"
 /clone_lib="NIH_MGC_90"
 /tissue_type="adenoecarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pCMV-Sport6; Site:1: NotI;
 Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 125 a 240 c 227 g 136 t
 ORIGIN

Query Match 49.3%; Score 677.4; DB 10; Length 728;
 Best Local Similarity 99.0%; Pred. No. 4.9e-120;
 Matches 724; Conservative 0; Mismatches 1; Indels 6; Gaps 4;

QY 192 cctggagcgtgcctgcctgcctctctgtgcgcgtgtgtcagtttgggagccgggacatc 251
 DB 1 CCGGCGGCTGGCTGCTGCGCTCTGCGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
 QY 252 gctgtccgc 308
 DB 61 GCTGTCCGCCACAG 120
 QY 309 gtccgaactgaatcccccagacagaagaagccagatcctcgccttctcctgaacccgact 368

Db 121 GTCCGAATCGAATCCCGAGACAGAAAGCCAGATCTCGCTTTCGACACCGACT 180
 Qy 369 agttcgcctctgcgaagtgtcactaaagccggaaaaacacgagctcgaagaagcatgac 428
 Db 181 AGTTGGGCTCTGCGAAGTGCACCTAAAGCGCGGAAACACGGGCTCCAAAGCGCATGCG 240
 Qy 429 agcccatatgaattcaccagcactgtgcagagcagcagcagcagcagcagcagcagcagc 488
 Db 241 AGCCATTATTAAGTTCTCATCCAGACCTGACAGAGACGACGACGACGACGAGGTGAGAGG 300
 Qy 489 gacagtagtgcctgcggaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 548
 Db 301 GACAGTAGTGGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
 Qy 549 ccaagatggggaagtttaagtcacccgggagcgtcgtcactactactgtactgtcagtgca 608
 Db 361 CCAAGATCGGGAGTTTAAGTACCCCGGGCTGGGCTTACTACCTGTACTGTACAGTGCA 420
 Qy 609 ctctgataagggaagagctgtctactaagcctgagcagcagcagcagcagcagcagcagcagc 668
 Db 421 CTTTGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
 Qy 669 cctgcgcctgcggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 728
 Db 481 CCTGGGCTGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
 Qy 729 cctctgcagagtgctgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 788
 Db 541 CTTCTGCGAGGAGTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 599
 Qy 789 cctccctgcggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 848
 Db 600 CCTCCCTCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 658
 Qy 849 tcaactgagggagcctgtgtctcccaagcagcagcagcagcagcagcagcagcagcagcagc 908
 Db 659 TCACGAGAGGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 717
 Qy 909 ctctgcgagcagc 919
 Db 718 CTCTGGGCGACC 728

RESULT 6
LOCUS BI966060 609 bp mRNA linear EST 23-OCT-2001
DEFINITION 1e72g04.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 cDNA 3' similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF
 APOPTOSIS ;, mRNA sequence.
ACCESSION BI966060
VERSION BI966060.1 GI:16340465
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 609)
AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
 Lemishka, I., Scaer, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistein, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
 M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T.,
 Jackson, Y. and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Juliana Brown
 (brownj@fas.harvard.edu)
 High quality sequence stop: 412.
 location/Qualifiers
 1. 609
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="Melton Normalized Human Islet 4 N4-HIS 1"
 /sex="Both"
 /tissue_type="Islets of Langerhans"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /note="Organ: Pancreas; Vector: pSPORT1; Site: 1; Not 1;
 Site_2: Sal 1; Starting library constructed using
 SuperScript Plasmid Library kit (Life Technologies). cDNA
 made by oligo-dT priming. Size-selected by column
 fractionation; average insert size 1.08 kb. Library was
 amplified once on solid support and plasmid DNA from
 library was prepared. The library DNA was normalized by
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 plasmid DNA was mixed with 5 micrograms PCR product
 representing library inserts and hybridized to an Ecot of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."

BASE COUNT 142 a 128 c 209 g 129 t 1 others
ORIGIN

Query Match 42.6%; Score 584.4; DB 10; Length 609;
 Best Local Similarity 99.7%; Pred. No. 3.2e-102;
 Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 787 accctccctgggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 846
 Db 609 ACCCTCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 550
 Qy 847 gtacactgagggagcctgtgtctcccaagcagcagcagcagcagcagcagcagcagcagcagc 906
 Db 549 GTTCACTGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 490
 Qy 907 ctctcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 966
 Db 489 CTCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 430
 Qy 967 cctcccttaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1026
 Db 429 CCTCCCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 370
 Qy 1027 ttccacctatcttaaacactcccccacagcagcagcagcagcagcagcagcagcagcagcagc 1086
 Db 369 TTCCACACTTATCTTAAACATCCCGACCGGACCTCCACCTACAGTACGCTCCCAAT 310
 Qy 1087 cctcagcccttgagggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1146
 Db 309 CCTACACCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 250
 Qy 1147 atgtgtcaactgtactctgtgtggaagagatgtgtggaagagagcagcagcagcagcagcagc 1206
 Db 249 ATTGTGTTCACTGTACTCTGTGTGGGAGAGATGGGTCCAGAGACCCCACTTAGGCACTA 190
 Qy 1207 agaagggcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1266
 Db 189 AGAGGGGCTGAGACTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 130
 Qy 1267 aatgtgagggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1326
 Db 129 AATGTGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 70

| FEATURES | source |
|------------|--|
| Db 867 | ctcccacaagctcgtcccaagctgtccgagctccctctgaagctctcttggaacccggtcc 926 |
| Db 240 | ctccccacagctgtgtccacagctgcccggctccctctgacagctcttggggacccgggtccc 299 |
| Qy 927 | ctctgcccaccctcagcgcgtctttgtctcagacctgcccctccctctagaagctgct 986 |
| Db 300 | ctttctcccccaccctctcagccgctcttttgctccagacctctccctctctagagctgacct 359 |
| Qy 987 | gggcctgtctaacggtgttttccatcccaataataacagatattccaccttatcttaaa 1046 |
| Db 360 | ggggctgtttcaccggtttttccatgccacatpaaatpcaatattccacattatcttaccaa 419 |
| Qy 1047 | ctcccccaccccccactctcaccctctactagtctcccccacatccctccttgaggcccc 1106 |
| Db 420 | cttccccacagcg-cacactctccacctctactagtctccccaattccctgacacctttgagagcccc 478 |
| Qy 1107 | cagtgatctcgactcccccttgccacacagaccccccaaggcatgtgtgtcactgtactctg 1166 |
| Db 479 | agtgatctactcgactccccccagtgagacacagaccgcccacagatagtgatctactctgt 538 |
| Qy 1167 | tgggaagaagtgatgggtccagaagaaccacactcaggcataagaagggtctgagacctggcg 1226 |
| Db 539 | gagggcaagagtgagggtccagaaagagc---cattgaggaaccacaaagaggggtgaacctgagagc 595 |
| Qy 1227 | caggaagccaagagagactg-ggctcaggccagagatcccccaatgtgaaggggcgaagaaac 1285 |
| Db 596 | agggaaagcaaaagagacactgagggccttaggcaagacattttcccaaatgacagggcgaagaaaa 655 |
| Qy 1286 | aagacaagctctctcccttgagaaatccctctgtgattttaa 1327 |
| Db 656 | caaggaagaccctcccatgagagatgccttgagatattgacaa 697 |
| RESULT 9 | |
| BI824443 | 531 bp mRNA linear EST 04-OCT-2001 |
| LOCUS | 603038693F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179510 5', |
| DEFINITION | mRNA sequence. |
| ACCESSION | BI824443 |
| VERSION | BI824443.1 GI:15935993 |
| KEYWORDS | EST. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| AUTHORS | Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo. |
| TITLE | 1 (bases 1 to 531) |
| JOURNAL | NIH-MGC http://mgc.nci.nih.gov/. |
| COMMENT | National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs@email.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
plate: LMNL1447 row: 3 column: 23
High quality sequence stop: 529. |
| | Location/Qualifiers |
| | 1..531 |

destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MC Library."

| Query Match | 37.6% | Score 516.8 | DB 10 | Length 531 |
|-----------------------|--|-------------------|--------------|------------------------------|
| Best Local Similarity | 99.4% | Pred. No. 2,9e-89 | | |
| Matches 529 | Conservative 0 | Mismatches 2 | Indels 1 | Gaps 1 |
| QY 661 | gtctgagccctgagctgcttgaagaaattctgaagcattcgagcagttccctcgggccc | | | 720 |
| Db 1 | GTGCTGGCCCTHGGCGCTGCTGGAGAAATTCACGACCTGGGGGAGATTCCTCGGGCCC | | | 60 |
| QY 721 | cagctcgagctctgccaagtgctgtgagctgttgccctcgagccagggctccctcg | | | 780 |
| Db 61 | CAGCTCCGCTCTGCGACAGGTGTCTGGGCTGTGGGCCCTCGGGCCA-GGTCCTCCCTGGG | | | 119 |
| QY 781 | atccgaccaccccccttggagcccatcctaaggctgcccccttcctaactctgagctc | | | 840 |
| Db 120 | ATCCGACCCCTTCCCTGGGGCCATCTCAAGGCTGCCCTCTCTCACTTCTGGACTC | | | 179 |
| QY 841 | tccagagttcaatgaggggacctggtctcccccacagtcgtcccaagctcgagctccct | | | 900 |
| Db 180 | TTCAGAGTTCACTGAGGGGGCCCTGGTCTCCCGCAGTGTGTCCAGAGCTGCCGCTCCCT | | | 239 |
| QY 901 | cgacagctctctgagcaccggtccctctgcccacccctaagccgtctctgtctcaga | | | 960 |
| Db 240 | CGACAGCTCTGTGGGACACCCGGTCCCTCTGGCCACCTCAGCCGCTTGTGCTCAGA | | | 299 |
| QY 961 | cctgcccctccctctagaagctgctgtgagctgttccaagttttccatcccaataat | | | 1020 |
| Db 300 | CTGTCCCTCTCTCTGAGAGGCTGCTGGGCCCTGTTCACGTGTTCATCCACATTAAT | | | 359 |
| QY 1021 | acagatcccaactcttcttaaacatccccacagcccaactctcaactcaactagctc | | | 1080 |
| Db 360 | ACAGATTCCTCCACTTTATCTTACACTCCCCACGCGCCACTCTCCACTACTAGCTC | | | 419 |
| QY 1081 | cccaatccctgaccctttagagcccccaagtgatctcgactcccccttggccaagaagcc | | | 1140 |
| Db 420 | CCCATCCCTGACCTTTGAGAGCCCCCAGATGATTCGACTCCCCCTGGCAGACAGACCC | | | 479 |
| QY 1141 | caaggcattgtgttactgtactctctgttggcaagaagtgtgtccagaagaagcc | | | 1192 |
| Db 480 | CAGGCAATTGTGTCTACTGTACTCTGTGGGCAAGATGGGTCCAGAGAAGCC | | | 531 |
| RESULT 10 | | | | |
| BF577781 | | 918 bp | mRNA | linear |
| LOCUS | 602092080F01 | | | EST 12-DEC-2000 |
| DEFINITION | | | Mus musculus | cDNA clone IMAGE:4206595 5', |
| ACCESSION | BF577781 | | | |
| VERSION | BF577781.1 | | | |
| KEYWORDS | EST. | | | |
| SOURCE | house mouse. | | | |
| ORGANISM | Mus musculus | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| AUTHORS | Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus. | | | |
| TITLE | 1 (bases 1 to 918) | | | |
| JOURNAL | NIH-MGC http://mgc.nci.nih.gov/. | | | |
| COMMENT | National Institutes of Health, Mammalian Gene Collection (MGC) | | | |
| | Unpublished (1999) | | | |
| | Contact: Robert Strausberg, Ph.D. | | | |
| | Email: cgabbs-remail.nih.gov | | | |
| | Tissue Procurement: Jeffrey E. Green, M.D. | | | |
| | cDNA Library Preparation: Life Technologies, Inc. | | | |
| | cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) | | | |
| | DNA Sequencing by: Incyte Genomics, Inc. | | | |
| | Clone distribution: MGC clone distribution information can be | | | |

FEATURES

CDS

ORIGIN

DB 24

03

ACCEPTED

| DB | QY | Db | RESULT 14 |
|---|-----|-----|-----------|
| 377 | 444 | 437 | BI966255 |
| <p> DEFINITION <i>Ie72g04.y1</i> Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens cDNA 5' similar to TR:043508 043508 TNF-RELATED WEAK INDUCER OF APOPTOSIS ;, mRNA sequence.
 ACCESSION BI966255
 VERSION BI966255.1 GI:16340660
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 REFERENCE Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS 1 (bases 1 to 456)
 Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scarce,M., Bresnelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarelisvilli,R., Williams,T., Jackson,Y. and Bowers,Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Other ESTs: ie72g04.x1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biolhp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Juliana Brown (brownjefas.harvard.edu)
 High quality sequence stop: 429.
 FEATURES
 source
 Location/Qualifiers
 1..456
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /sex="Both"
 /tissue_type="Islets of Langerhans"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1; Site_2: Sal 1; Starting library constructed using Superscript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoT of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."
 BASE COUNT 57 a 188 c 105 g 106 t
 ORIGIN
 Query Match 32.1%; Score 440.2; DB 10; Length 456;
 Best Local Similarity 99.1%; Pred.No. 1.4e-74; </p> | | | |

| | | | | | | | | | |
|---------|------|--|------|------------|----|--------|----|------|----|
| Matches | 453; | Conservative | 0; | Mismatches | 3; | Indels | 1; | Gaps | 1; |
| OY | 624 | gagctgtcactgaagctggagcttgccttgtagatggtgtgtcgttcgtccctcgctcctctga | 683 | | | | | | |
| Dd | 1 | GGCTGTCAACCGAAGCTGGACTTCTGATGAT-GTATGGCGCCGCGCCTGCCGTGA | 59 | | | | | | |
| OY | 684 | gaaattccaagcaactggcgacagtccctcctggagcccagaacctcgacctccttcgaagtgc | 743 | | | | | | |
| Dd | 60 | GGAATTTCACGCACATGGCGGAGTTCCCTCGGGCCCCAGCTCCGCTCTGCCAGGTTC | 119 | | | | | | |
| OY | 744 | tggagctgttgacctgagcgagaggltccctccctcgatlcgcgacaacctcccttcggacca | 803 | | | | | | |
| Dd | 120 | TGGGCTGTTGGCCATCGCGGCCAGGGGTCCMCCCTGGGGAATCGCACACTCCCTGGGGCCA | 179 | | | | | | |
| OY | 804 | tctcaagctgcacccttcctaacttggagctcttcgaagttcaactgaaggccct | 863 | | | | | | |
| Dd | 180 | TCTCAAGGTGGCCCTCTCTCATCTACTTGCGACTTCCAGGTTCACTGAGGGGCCCT | 239 | | | | | | |
| OY | 864 | ggtctcccccaagttgtcccaagctgagcggtcccttcctgaagctcttggcacccggt | 923 | | | | | | |
| Dd | 240 | GGTCTCCCGCAGTGTCTCCAGGCTGGCGGCTCCCTCGACAGCTCTGGGACCCTG | 299 | | | | | | |
| OY | 924 | cccccttgcaccacacctgaagcgcctcttgtgtccagacctgcccctcctataagctg | 983 | | | | | | |
| Dd | 300 | CCCCTGTGCCCCACGCTCAGCGCGCTTGTGCTCCAGACTGCGCCTCCCTCAAGGGCTG | 359 | | | | | | |
| OY | 984 | cctggcgctgtacagtggtttccalcacacaaataaacaglatcccaactctta | 1043 | | | | | | |
| Dd | 360 | CCTGGGCGCTGTTCAGTGTTCATCCACATAAATACAGTAATCCACACTTATCTTA | 419 | | | | | | |
| OY | 1044 | caactcccccaagcgcaactctcaccctcacactc | 1080 | | | | | | |
| Dd | 420 | CAACTCCCCACCGCCACTCTCACCTCACAATCACTG | 456 | | | | | | |

RESULT 15
BM128059

```
LOCUS      BM128059              440 bp          mRNA           linear       EST 27-NOV-2001
DEFINITION 1f09b06.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
            CDNA 5' similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF
            APOPTOSIS ; , mRNA sequence.
ACCESSION   BM128059
VERSION     BM128059.1 GI:17122611
KEYWORDS    EST.


ORGANISM   Homo sapiens
            human.



REFERENCE 1 (bases 1 to 440)
            Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemshaha,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blisstein,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
            ,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T.,
            Jackson,Y. and Bowers,Y.
            Endocrine Pancreas Consortium
            Unpublished (2000)
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu



JOURNAL    Library was constructed by Dr. Douglas Melton DNA sequencing by:
COMMENT     Washington University Genome Sequencing Center for information on
            obtaining a clone please contact: Juliana Brown
            (brownefas.harvard.edu)



FEATURES   High quality sequence stop: 415.
            Location/Qualifiers
                1..440
                    /organism="Homo sapiens"


```

```
/db_xref="taxon:9606"  
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"  
/sex="Both"  
/tissue_type="Islets of Langerhans"  
/dev_stage="Adult"  
/lab_host="DH10B"  
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;  
Site_2: Sal 1; Starting library constructed using  
SuperScript Plasmid Library kit (Life Technologies). cDNA  
made by oligo-dT priming. Size-selected by column  
fractionation; average insert size 1.08 kb. Library was  
amplified once on solid support and plasmid DNA from  
library was prepared. The library DNA was normalized by  
method #4 from Bonaldo, Lennon, and Soares 1996 Genome  
Research 6:791-806; 0.5 microgram single-stranded library  
plasmid DNA was mixed with 5 micrograms PCR product  
representing library inserts and hybridized to an EcoT of  
20. Single-stranded (unhybridized) plasmids were isolated  
by hydroxyapatite chromatography and used to make this  
library."
```

```
BASE COUNT      71 a      186 c      86 g      97 t  
ORIGIN
```

```
Query Match      31.8%; Score 436.8; DB 10; Length 440;  
Best Local Similarity 99.5%; Pred. No. 6.3e-74;  
Matches 438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 771 ctccctgagatccgcacccctccctgggcccattccaaagctgccccctctctaccta 830  
Db 1 ctccctgcccagatccgcacccctccctgggcccattccaaagctgccccctctctaccta 60  
  
QY 831 ctctggactcttcaggcttcaactgagggccctgctctcccaagtcgtcccaagctgc 890  
Db 61 ctctggactcttccaggttcaactgagggccctgctctcccaagtcgtcccaagctgc 120  
  
QY 891 cggctccctcgacaagctctctgagcaccgcgctccctctgcccacacctagcgctct 950  
Db 121 cggctccctcgacaagctctctgagcaccgcgctccctctgcccacacctagcgctct 180  
  
QY 951 ttgctcagagctgccccctccctctagagctgctgggctgttcaagtgtttccatc 1010  
Db 181 ttgctcagagctgccccctccctctagagctgctgggctgttcaagtgtttccatc 240  
  
QY 1011 ccacataaatacagla ttcccaacttlatcttaaacatcccccacgcgccaactctcacc 1070  
Db 241 ccacataaatacagla ttcccaacttlatcttaaacatcccccacgcgccaactctcacc 300  
  
QY 1071 tcaactagctccccaatccctgaccccttgagcccccaagtgatctgaactccccctgac 1130  
Db 301 tcaactagctccccaatccctgaccccttgagcccccaagtgatctgaactccccctgac 360  
  
QY 1131 caaagaccccccaaggaatggtttcaactgtacctgtgggcaagaagtgttccagaagac 1190  
Db 361 caaagaccccccaaggaatggtttcaactgtacctgtgggcaagaagtgttccagaagac 420  
  
QY 1191 cccacttcaggcaactaagag 1210  
Db 421 cccacttcaggcaactaagag 440
```

Search completed: June 22, 2002, 10:49:15
Job time: 18070 sec